



Língua eletrónica potenciométrica: uma ferramenta para análise da qualidade do mel.

Mara E.B.C. de Sousa Dias

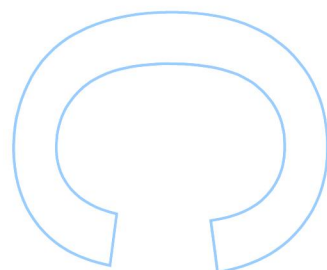
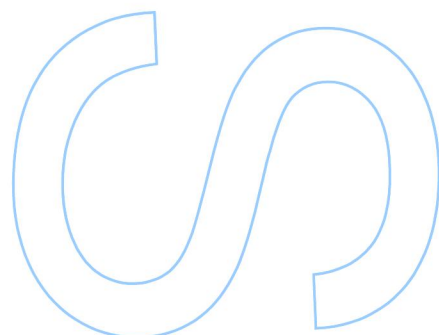
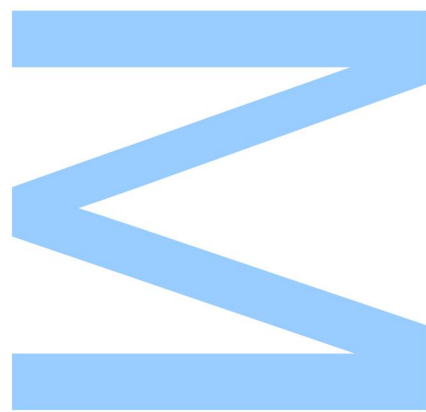
Mestrado em Química
Departamento de Química e Bioquímica
2013

Orientador

Adélio A.S.C. Machado, Professor Catedrático, Faculdade de Ciências

Coorientador

António M. Peres, Professor Adjunto, Instituto Politécnico de Bragança

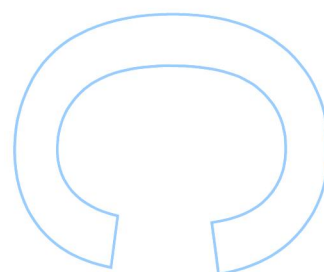
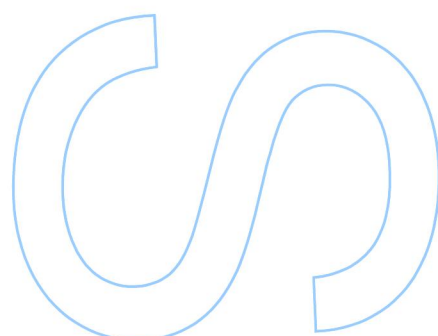
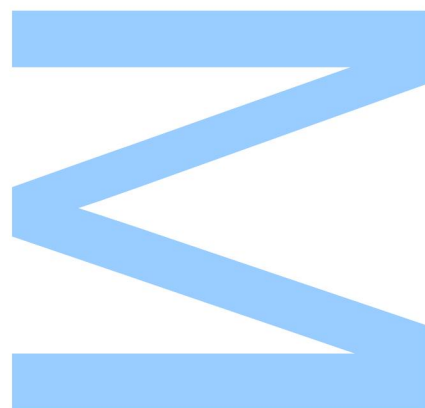




Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____/____/____



Ao meu marido e filhos

Agradecimentos

Ao Professor Doutor Adélio Machado, orientador científico, pela disponibilidade, apoio e valiosas sugestões ao longo de todo o trabalho.

Ao Professor Doutor António Peres, co-orientador científico, pelas críticas oportunas.

À Professora Doutora Letícia Estevinho, pela disponibilidade e colaboração nas análises polínicas e de cor das amostras de mel estudadas.

Ao Engº Jorge Sá Morais, pelo apoio na realização das análises por cromatografia líquida (HPLC).

À Escola Superior Agrária do Instituto Politécnico de Bragança, pela oportunidade oferecida para a realização deste trabalho no Laboratório de Química e Bioquímica Aplicada.

Resumo

Neste trabalho investigou-se a possibilidade de discriminação de diferentes tipos de méis monoflorais com uma língua eletrónica potenciométrica construída pelo método *print-screen* e constituída por sensores de sensibilidade cruzada. Usaram-se 89 amostras de mel, provenientes de todas as regiões apícolas de Portugal continental relativas a 3 anos apícolas (2009-2011), que foram analisadas quanto à cor, perfil polínico e conteúdo de glucose e frutose. As mesmas amostras foram analisadas com recurso à língua eletrónica, tendo sido registados os seus perfis potenciométricos. Constatou-se a existência de uma acentuada variabilidade das amostras com base no perfil de cores obtido, desde o extra-claro até ao escuro; pelos níveis de concentração de frutose que variaram entre 33 e 57 g/100 g de mel e os de glucose entre 20 e 37 g/100 g de mel; assim como quanto ao tipo e níveis de pólen predominantes detectados. Estes dados permitiram classificar 75 amostras como pertencendo a 8 tipos de mel monofloral: *Lavandula sp.*, *Echium sp.*, *Rubus sp.*, *Erica sp.*, *Prunus sp.*, *Castanea sp.*, *Trifolium sp.* e *Eucalyptus sp.* (os dois últimos, apenas com 2 méis presentes na amostragem realizada).

Atendendo à redundância da informação contida nos perfis de sinais potenciométricos da língua eletrónica, foram testados três algoritmos meta-heurísticos de seleção de variáveis (arrefecimento simulado, genético e melhoria local), em paralelo com o método de análise discriminante linear, para estabelecer modelos de classificação para o tipo de mel. Devido ao elevado número de tipos de mel presentes nas amostras foi necessário definir uma metodologia de tratamento de dados que permitisse efetuar de forma robusta, com a língua eletrónica potenciométrica, a classificação das amostras no respetivo grupo de mel monofloral. Os méis monoflorais foram divididos em 3 grupos de acordo com o perfil de cores obtido: branco (incluindo os méis extra-branco e branco), âmbar (incluindo os méis âmbar extra claro, âmbar claro e âmbar) e escuro (constituído pelos méis escuros). Este procedimento, associado aos sub-conjuntos de sensores selecionados com o algoritmo de Arrefecimento Simulado (algoritmo que apresentou sempre os melhores resultados), permitiu obter um modelo de discriminação linear mais simples (com menor número de sensores no modelo de análise discriminante linear) e com critério de qualidade do ajuste elevado e, principalmente, com 100% de classificações corretas para o procedimento de validação cruzada *leave-one-out*.

Para as amostras classificadas de cor branco obteve-se um modelo de análise discriminante linear com 6 sensores que permitiu obter 100% de classificações corretas, na validação cruzada *leave-one-out*, entre as amostras dos dois grupos de méis monoflorais de *Echium* e *Lavandula*. Para as amostras de cor âmbar, o melhor modelo continha 16 sensores e permitiu discriminar corretamente os 4 tipos de mel monofloral: *Echium*,

Lavandula, *Prunus* e *Rubus*. No grupo de cor escura, o modelo selecionado possuía 7 sensores e permitiu discriminar corretamente os 3 grupos de mel monofloral: *Castanea*, *Erica* e *Rubus*.

O desempenho analítico observado pela língua eletrónica potenciométrica construída sugere que esta técnica poderá ser uma ferramenta de auxílio na análise rotineira de classificação dos méis, permitindo às empresas, que produzem e comercializam o mel, ter a capacidade de classificar/confirmar se uma amostra de mel é monofloral e de que tipo, com a mais-valia económica que essa classificação aporta por comparação com um mel multifloral, garantindo também a autenticidade do produto embalado, de um ponto de vista floral.

Palavras chave: Língua eletrónica; méis monoflorais; análise discriminante linear; seleção de sensores; algoritmos meta-heurísticos; algoritmo de arrefecimento simulado; algoritmo genético; algoritmo de melhoramento local restringido

Abstract

The performance of a potentiometric electronic tongue constructed by the print-screen method and comprising cross-sensitivity sensors, for discriminate different types of monofloral honeys was evaluated. In total, 89 samples of different honeys obtained all across continental Portugal, were used and analysed concerning color, pollen profile, glucose and fructose contents. All samples were also analysed using the electronic tongue and the potentiometric signal profiles were recorded. The honey samples showed a high variability that was demonstrated by their different colors, yielding from the extra-light to the dark colors; by the different fructose (ranging from 33 to 57 g/100 g of honey) and glucose (between 20 and 37 g/100 g honey) concentration levels; as well as by the type and levels of the major pollens identified. Based on the pollen analysis 75 samples were classified as monofloral honeys: *Lavandula* sp., *Echium* sp., *Rubus* sp., *Erica* sp., *Prunus* sp., *Castanea* sp., *Trifolium* sp. and *Eucalyptus* sp. (although the last two with only two samples).

The potentiometric signal profiles recorded showed a high level of redundancy. So, three variable selection meta-heuristic algorithms were applied (simulated annealing, genetic and restricted local improvement) to establish the most informative predictors (sensors) when using in simultaneous a linear discriminant analysis method to establish classification models according to the type of monofloral honey. Due to the high variability of monofloral honey in the sampling database, a methodology was chosen to allow using the electronic tongue for monofloral honey classification.

Samples classified as monofloral honeys were split into 3 groups according to their colors: white (including extra-white and white honeys), amber (including extra light amber, light amber and amber honeys) and dark (composed of dark honeys). This procedure together the subsets of sensors selected using the Simulated Annealing algorithm (algorithm that allowed the establishment of the best classification model), allowed a simple linear discrimination model (with a reduced number of sensors in the model) with high values for the goodness of fit criterion, and especially with 100% of correct classifications for the leave-one-out cross-validation procedure.

For samples classified as white honeys, a linear discriminant analysis model with 6 sensors allowed a 100% of correct monofloral honey classifications as *Echium* or *Lavandula*, using a leave-one-out cross-validation procedure. For samples classified as amber, the best model included the signals of 16 sensors and allowed a full discrimination between the samples from the 4 types of monofloral honeys: *Echium*, *Lavandula*, *Prunus* and *Rubus*. Finally, for the dark color group, the selected model had 7 sensors and allowed to discriminating correctly the 3 groups of monofloral honeys: *Castanea*, *Erica* and *Rubus*.

The results achieved with the proposed potentiometric electronic tongue suggest that this analytical technique may be an important tool to assist in the routine honey classification, allowing companies that produce and sell honey, to have the capability to classify/confirm whether a honey is monofloral and of which kind, with the economic added value that it brings in relation to multifloral classified honeys, ensuring also the authenticity of the packaged product, from a floral point of view.

Keywords: Electronic tongue; Monofloral honeys; Linear discriminant analysis; Sensors selection; Meta-heuristic algorithms; Simulated annealing algorithm; Genetic algorithm; Restricted local improvement algorithm.

Índice geral

| | |
|--|------|
| Resumo | iii |
| Abstract | v |
| Índice geral | vii |
| Índice de tabelas | ix |
| Índice de figuras | x |
| Lista de abreviaturas | xi |
| PREÂMBULO | xiii |
| a. Enquadramento | xiii |
| b. Objetivos | xiii |
| c. Estrutura do trabalho | xiv |
| 1. INTRODUÇÃO | 1 |
| 1.1. O mel | 1 |
| 1.1.1. <i>Produção do mel pela abelha</i> | 1 |
| 1.1.2. <i>Composição química do mel</i> | 1 |
| 1.1.3. <i>Pólen no mel</i> | 2 |
| 1.1.4. <i>Classificação do mel</i> | 3 |
| 1.1.5. <i>O mel em Portugal</i> | 4 |
| 1.1.6. <i>Controlo de qualidade</i> | 5 |
| 1.2. Língua eletrónica | 6 |
| 1.2.1. <i>Sensores químicos</i> | 6 |
| 1.2.2. <i>Processamento de sinais</i> | 7 |
| 1.2.3. <i>Exemplos de aplicação de LEs</i> | 11 |
| 1.3. Génese do trabalho | 12 |
| 2. Material e métodos | 13 |
| 2.1. Amostragem | 13 |
| 2.2. Caracterização da cor do mel | 13 |
| 2.3. Análise polínica | 14 |
| 2.4. Análise por HPLC | 15 |
| 2.4.1. <i>Equipamento</i> | 15 |
| 2.4.2. <i>Preparação de soluções</i> | 15 |

| | |
|---|----|
| 2.5. Análise com a língua eletrónica | 16 |
| 2.5.1. Montagem do sistema de multi-sensores..... | 16 |
| 2.5.2. Sensores químicos..... | 17 |
| 2.5.3. Equipamento | 17 |
| 2.5.4. Análise com a língua eletrónica..... | 18 |
| 2.6. Análise de dados..... | 19 |
| 3. Resultados | 21 |
| 3.1. Cor do mel | 21 |
| 3.2. Análise polínica..... | 23 |
| 3.3. Análise por HPLC..... | 27 |
| 3.4. Análise com a língua eletrónica | 32 |
| 3.4.1. Classificação de amostras de cor branca..... | 38 |
| 3.4.2. Classificação de amostras de cor âmbar..... | 40 |
| 3.4.3. Classificação de amostras de cor escura..... | 43 |
| 4. Conclusões..... | 47 |
| Referências | 49 |

Índice de tabelas

| | |
|--|----|
| Tabela 1 – Exemplo de alguns resultados analíticos referentes a análises polínicas e químicas de méis monoflorais de <i>Castanea</i> obtidos em vários trabalhos [13]. | 4 |
| Tabela 2 – Cores padrão do mel designadas pela USDA e respetivos intervalos de cor da escala mm PFund. | 14 |
| Tabela 3 – Composição das membranas poliméricas. | 18 |
| Tabela 4 – Número de amostras com cores padrão do mel designadas pela USDA. | 21 |
| Tabela 5 – Percentagem de méis com cada tipo de pólen presente e o valor mínimo e máximo da percentagem de cada pólen nas amostras de mel. | 24 |
| Tabela 6 – Exemplo de planta associada a cada género identificado através da análise polínica. | 25 |
| Tabela 7 – Classificação e os 3 pólenes mais predominantes das amostras. | 26 |
| Tabela 8 – Valores globais da classificação e da cor das amostras monofloral/multifloral. | 27 |
| Tabela 9 – Intervalo de concentrações e parâmetros da calibração para a análise da glucose e frutose por HPLC. | 29 |
| Tabela 10 – Resultados do estudo da repetibilidade, da precisão intermédia e da exatidão com soluções de controlo de qualidade. | 30 |
| Tabela 11 – Resultados do estudo da repetibilidade e da precisão intermédia de 3 amostras. | 30 |
| Tabela 12 – Concentrações de glucose e frutose nas amostras de mel analisadas por HPLC | 31 |
| Tabela 13 – Seleção de variáveis independentes através de 3 algoritmos para obter modelos de discriminação entre 2 grupos de mel monofloral de cor branco. | 39 |
| Tabela 14 – Seleção de variáveis independentes através de 3 algoritmos para obter modelos de discriminação entre 4 grupos de mel monofloral de cor âmbar. | 42 |
| Tabela 15 – Seleção de variáveis independentes através de 3 algoritmos para obter modelos de discriminação entre 3 grupos de mel monofloral de cor escura. | 44 |

Índice de figuras

| | |
|--|----|
| Figura 1 – Escala impressa de cores padrão do mel e respetivos intervalos de valores mm PFund..... | 14 |
| Figura 2 – Sistema de multi-sensores..... | 17 |
| Figura 3 – Montagem para a análise com a língua eletrónica..... | 19 |
| Figura 4 – Amostras de mel usadas neste trabalho de acordo com a escala de cores padrão do mel impressa e as definidas pela USDA..... | 22 |
| Figura 5 – Distribuição da cor Pfund (mm) e ajuste da função de distribuição log-normal. | 23 |
| Figura 6 – Cromatogramas da turanose, trehalose, maltose e de uma mostra de mel. | 28 |
| Figura 7 – Perfil dos sinais obtidos da análise das amostras com a língua eletrónica. | 32 |
| Figura 8 – Extremos quartis do potencial (mV) em função dos sensores. | 33 |
| Figura 9 – Esquema da metodologia de tratamento de dados usada para estabelecer o melhor modelo ADL..... | 35 |
| Figura 10 – Análise discriminante linear entre 8 diferentes tipos de mel. | 36 |
| Figura 11 – Análise discriminante linear entre 6 diferentes tipos de mel. | 37 |
| Figura 12 – Análise discriminante linear entre os 3 grupos de cor associados à seleção primária das amostras: 1 – branco; 2 – âmbar; 3 – escuro..... | 38 |
| Figura 13 – Distribuição de frequências dos dados referentes à primeira função discriminante da classificação de dois grupos de méis brancos monoflorais: <i>Echium</i> (Ech) e <i>Lavandula</i> (Lav). | 41 |
| Figura 14 – Análise discriminante linear de 4 grupos de méis monoflorais de cor âmbar: <i>Echium</i> (Ech), <i>Lavandula</i> (Lav), <i>Prunus</i> (Pru) e <i>Rubus</i> (Rub). | 43 |
| Figura 15 – Análise discriminante linear de 3 grupos de méis monoflorais de cor escura: <i>Castanea</i> (Cas), <i>Erica</i> (Eri) e <i>Rubus</i> (Rub). | 45 |

Lista de abreviaturas

A – Âmbar

AC – Âmbar claro

Aca – *Acacia sp.*

ACP – Análise de Componentes Principais

ADL – Análise Discriminante Linear

AEC – Âmbar extra claro

AS – Arrefecimento Simulado

B – Branco

Cas – *Castanea sp.*

ccr12 – coeficiente da primeira raiz de Roy (*coefficient and Roy's first root statistic, ccr₁²*)

Cre – *Crepis sp.*

Cyt – *Cytisus sp.*

E – Escuro

EB – Extra-branco

Ech – *Echium sp.*

Eri – *Erica sp.*

Euc – *Eucalyptus sp.*

FNAP – Federação Nacional de Apicultores de Portugal

Foe – *Foeniculum sp.*

GEN – Genético

Gen – *Genista sp.*

Hel – *Helianthus sp.*

HPLC – Cromatografia líquida de alto desempenho (abreviatura do inglês, *high performance liquid chromatography*)

IHC – Comissão Internacional do mel (abreviatura do inglês, *International Honey Commission*)

IR – Índice de Refração

IUPAC – União Internacional da Química Pura e Aplicada (abreviatura do inglês, *International Union of Pure and Applied Chemistry*)

Lav – *Lavandula sp.*

LE – Língua Eletrónica

Leo – *Leontodon sp.*

LOO – “deixar um de fora” (abreviatura do inglês, *leave-one-out*)

Med – *Medicago sp.*

Men – *Mentha sp.*

Mim – *Mimosaceae sp.*

MLR – Melhoramento Local Restringido

MQP – Mínimos Quadrados Parciais

MQP-AD – Mínimos Quadrados Parciais na Análise Discriminante

Per – *Persea sp.*

Pin – *Pinus sp.*

Pru – *Prunus sp.*

PVC – Cloreto de polivinilo (abreviatura do inglês, *Polyvinyl Chloride*)

RLM – Regressão Linear Múltipla

RNA – Redes Neurais Artificiais

Rub – *Rubus sp.*

San – *Sandix sp.*

Thy – *Thymus sp.*

Til – *Tilia sp.*

Tri – *Trifolium sp.*

USDA – Departamento de Agricultura dos Estados Unidos (abreviatura do inglês, *United States Department of Agriculture*)

UV-Vis – Ultravioleta-Visível

PREÂMBULO

a. Enquadramento

As empresas que produzem e comercializam o mel têm a necessidade de classificar os diferentes tipos de méis, uma vez que o mel monofloral tem um valor económico mais elevado do que o multifloral. A classificação dos méis em méis monoflorais ou multiflorais é vulgarmente realizada através de uma análise polínica, que é demorada e exige um técnico especializado na identificação dos diferentes tipos de pólen. Quando o perfil polínico não permite uma classificação correta, as amostras são sujeitas a uma análise sensorial para confirmação, mas esta é também uma análise especializada e limitada quanto ao número de amostras que é possível analisar por dia. Assim, há a necessidade de encontrar novas ferramentas analíticas que tornem mais fácil a classificação do mel no processo industrial, a fim de garantir a autenticidade dos seus produtos.

A língua eletrónica potenciométrica é uma recente aplicação analítica cujas potencialidades na área alimentar são vastas, pois permite mimetizar a análise sensorial quer ao nível qualitativo quer ao nível quantitativo. É uma ferramenta de trabalho que permite a classificação, discriminação e quantificação, aplicando métodos estatísticos multivariados não supervisionados e supervisionados aos perfis de sinais de potencial referentes a parâmetros químicos presentes nas amostras líquidas analisadas.

Foi neste âmbito que se decidiu investigar se a língua eletrónica potenciométrica poderia constituir uma metodologia analítica rápida para distinguir méis monoflorais, permitindo reduzir o número de amostras que necessitariam de ser sujeitas a análise polínica e sensorial, que seriam só as de classificação duvidosa.

b. Objetivos

O **objetivo geral** deste trabalho consistiu em aplicar a língua eletrónica potenciométrica na análise de méis, com vista a estabelecer a metodologia estatística que permitisse discriminar corretamente as amostras em grupos de méis monoflorais.

Os **objetivos específicos** foram os seguintes:

- obter um elevado número de amostras de mel;
- construir a língua eletrónica pelo método *print-screen* com membranas poliméricas de policloreto de vinilo baseadas na combinação de 5 plastificantes e 5 compostos aditivos;
- efetuar análise polínica, da cor e dos açúcares maioritários (glucose e frutose por cromatografia líquida de alto desempenho) em todas as amostras;
- efetuar a análise destas com a língua eletrónica;

- para tratamento de dados, usar técnicas de seleção de variáveis (sensores), tais como, algoritmos meta-heurísticos (arrefecimento simulado, genético e melhoramento local restringido), aliadas ao método multivariado de classificação análise discriminante linear;
- definir a melhor metodologia de tratamento de dados para garantir classificações corretas do grupo de mel monofloral de cada amostra com a língua eletrónica.

c. Estrutura do trabalho

O presente trabalho inicia-se com o capítulo de **Introdução**, resultante de uma pesquisa bibliográfica, constituída por uma breve descrição sobre o que é o mel e a sua composição química, bem como sobre a sua caracterização por análise polínica, fazendo menção à sua produção em Portugal e ao seu controlo de qualidade. Neste capítulo também se faz referência ao estado da arte das línguas eletrónicas, bem como a estudos sobre a sua aplicação ao mel.

No segundo capítulo, **Material e métodos**, faz-se a descrição da amostragem realizada e descrevem-se as metodologias aplicadas na caracterização da cor e análise polínica do mel. Em relação à análise por cromatografia líquida de alto desempenho (HPLC), apresenta-se o equipamento e preparação de soluções; para a análise com a língua eletrónica, descreve-se o sistema de multi-sensores, os sensores químicos usados e o equipamento. Abordam-se também os métodos estatísticos multivariados aplicados aos dados analíticos que foram usados no capítulo seguinte.

No terceiro capítulo, **Resultados e discussão**, apresentam-se os tratamentos de resultados obtidos das análises de cor, perfil polínico, HPLC e língua eletrónica, bem como a discussão dos resultados referentes aos tratamentos estatísticos aplicados.

No quarto capítulo, **Conclusões finais**, compilam-se as conclusões gerais mais relevantes obtidas no capítulo anterior e faz-se também uma apreciação global dos resultados tendo em vista os objetivos.

Por fim, apresenta-se a **Bibliografia** utilizada.

1. INTRODUÇÃO

1.1. O mel

O mel é uma substância doce produzida por abelhas a partir do néctar de flores, de secreções doces de outras partes da planta e de excreções dos insetos sugadores de plantas [1].

1.1.1. *Produção do mel pela abelha*

O produto mais comum é o produzido pelas abelhas melíferas (género *Apis*). A produção do mel é iniciada com um primeiro tratamento das matérias recolhidas, na boca e “estômago do mel” da abelha, usando substâncias específicas próprias – enzimas. Estas permitem quebrar os açúcares complexos (principalmente a sacarose) em açúcares mais simples (glucose e frutose), que são mais facilmente digeríveis e permitem uma maior conservação do mel. Na colmeia, as abelhas coletoras (campeiras) regurgitam o néctar e passam-no para as abelhas operárias da colmeia que continuam com o processo de transformação do néctar. Após este tratamento, o produto é colocado nos alvéolos de cera e, gradualmente, é transformado em mel por evaporação da água. O mel fica armazenado na colmeia para amadurecer e é usado como fonte primária de alimento para o inverno [2,3].

Como ao longo de um ano apícola o período de florescimento é diferente de planta para planta, o mel é uma mistura de néctares e por isso tem uma composição variável dependendo das fontes florais [3,4]. Outros factores que afectam a composição do mel, indiretamente ou diretamente, são o clima e o processamento do mel [4].

1.1.2. *Composição química do mel*

Globalmente, apresentando valores médios, o mel é constituído principalmente por açúcares (aproximadamente, 80% da sua massa), nomeadamente glucose (31,3%), frutose (38,2%), sacarose (0,7%), outros dissacarídeos (5%) e polissacarídeos (3,1%). Apresenta quantidades de água geralmente inferiores a 20% (no néctar, 80%) e valores de pH entre 3,5 e 4,5. O mel também contém uma grande variedade de outras substâncias em pequenas concentrações (<2%) como, por exemplo, aminoácidos e proteínas, sais minerais, compostos fenólicos (ácidos fenólicos e flavonoides), vitaminas, etc [2,3,5].

A composição química do mel afeta as suas propriedades físicas (índice de refração, viscosidade, densidade e calor específico), bem como o aroma e sabor. Estes estão associados aos componentes ácidos orgânicos e outros voláteis (álcoois, cetonas, aldeídos e ésteres) presentes no mel [2,3]. A cor resulta da mistura de vários méis presentes na colmeia, devido à recolha de néctar em diferentes florações ao longo do ano apícola, e está

diretamente dependente da composição em sais minerais, da fonte de néctar, de níveis elevados de aminoácidos e da presença de compostos fenólicos [2].

1.1.3. Pólen no mel

Outro produto que as abelhas obtêm das flores é o pólen, rico em proteínas (principal fonte de alimentação das abelhas), substâncias lipídicas, sais minerais e vitaminas [6,7]. É o alimento essencial para o crescimento de larvas de abelhas e jovens adultos [6]. A colheita do pólen não é feita na mesma altura do dia da recolha do néctar, pois a produção de pólen ou néctar na maioria das flores só acontece a determinadas horas, não coincidentes [7].

Por isso, a presença de pólen no mel deve-se a três possíveis situações, associadas à ação das abelhas [8]:

- os grãos de pólen da flor são desalojados das anteras e caem no néctar que é sugado pelas abelhas e armazenado no seu estômago;
- os grãos de pólen ficam presos aos cabelos, pernas, antenas e até mesmo nos olhos das abelhas e, mais tarde, quando o néctar é regurgitado, parte é depositado em células abertas dos favos de colmeia;
- as abelhas coletoras de mel agitam o seu corpo para remover os pólenes e durante este processo, o pólen cai em alvéolos abertos.

O teor de pólen no mel está relacionado com as características morfológicas das flores e pólen e é relativamente constante para cada uma das espécies individuais, de modo que é possível avaliar de forma precisa qual a abundância de cada pólen no néctar. Se estas características levam a uma elevada abundância de pólen no néctar, então o pólen é classificado de sobre-representado; se a abundância é baixa, o pólen é considerado sub-representado; nos casos intermédios, o pólen é considerado como tendo uma representatividade normal [9,10].

Como resultado da presença de pólen no mel verifica-se que é possível identificar a origem das fontes florais (origem botânica) e geográficas do mel produzido pelas abelhas [8-10] através de uma análise polínica. Esta corresponde à contagem dos grãos de pólen (baseada na frequência relativa dos diversos tipos de pólen) presentes no mel, segundo o género [9,10].

Esta análise polínica, combinada com critérios físico-químicos e características organoléticas, é usada para controlo de qualidade, pois permite detetar misturas de méis de diferentes locais e dar informações gerais sobre a área de recolha (clima e a condição geográfica, botânica e agrícola) [11]. Esta informação é importante uma vez que o mel tem cor, sabor e textura de acordo com as fontes de néctares.

Considerando os pólenes predominantes e a sua respectiva classificação em relação aos níveis de representação, o mel pode ser classificado em mel monofloral de uma determinada planta (origem botânica) ou multifloral. A introdução desta informação no rótulo do produto é também essencial por afetar o preço do produto no mercado, pois os méis monoflorais têm maior valor comercial.

1.1.4. Classificação do mel

Em geral, considera-se um mel monofloral de uma dada espécie vegetal quando o pólen dessa espécie é predominante (superior a 45%, representatividade normal) em relação ao total dos pólenes observados [10,12]. Mas também há espécies vegetais com sub-representação em pólen, ou seja, a quantidade de néctar retirado da planta não está de acordo com a quantidade de pólen representado no mel. Um exemplo é o pólen de *Lavandula sp.*, cuja percentagem de predominância do pólen a partir da qual se pode considerar esse mel monofloral de *Lavandula sp.* é de 15% (pólen sub-representado) [12]. Situação inversa é a sobre-representação polínica de espécies vegetais como o pólen de *Castanea sp.*, cujo mel monofloral deve ter uma predominância deste pólen a partir dos 90% (pólen sobre-representado) [12]. No mel de *Eucalyptus sp.*, a predominância do pólen de *Eucalyptus sp.* deve ser superior a 70% [12]. Como regra, a classificação de mel monofloral de *Lavandula* prevalece em caso de haver outro pólen predominante.

Para uma correta classificação de um mel é recomendado que se obtenham outras características, como a sensorial e físico-químicas que, em conjunto, permitem ter uma maior confiança na determinação e controlo da origem botânica dos méis [10,13]. Como a variabilidade natural do mel não permite definir limites precisos de predominância do pólen para um dado tipo de mel (que, em geral, resulta de uma mistura de vários néctares), a interpretação depende da experiência dos analistas na avaliação dos resultados globais e do seu conhecimento do produto [10]. Aliás, quando há dificuldades em estabelecer a origem biológica, a decisão final é definida com base na avaliação organoléptica de um perito [14].

Um estudo de revisão bibliográfica sobre as características gerais de alguns méis monoflorais Europeus (15 tipos de mel monofloral), referentes a análises polínicas e físico-químicas, mostra a variabilidade nos méis de diferentes países classificados como sendo monofloral do mesmo tipo de pólen [13]. Como exemplo desta possível variabilidade, na Tabela 1 apresentam-se alguns dos dados de caracterização físico-química e polínica de méis monoflorais de *Castanea* compilados nesse estudo [13].

O mel monofloral Português tem sido alvo de vários estudos, como por exemplo, ao nível da caracterização polínica, físico-química, sais minerais, microbiológica e de atividade

biológica. Alguns exemplos de méis monoflorais Portugueses usados nestes estudos são: *Erica sp.* da região de Trás-os-Montes [15]; *Eucalyptus sp.*, *Lavandula sp.*, *Erica sp.*, *Cytisus sp.* da Serra da Lousã [16]; *Erica sp.*, *Eucalyptus sp.*, *Citrus sp.*, *Lavandula sp.* e *Echium sp.* provenientes de várias regiões de Portugal [17].

Tabela 1 – Exemplo de alguns resultados analíticos referentes a análises polínicas e químicas de méis monoflorais de *Castanea* obtidos em vários trabalhos [13].

| Parâmetro | Dados analíticos provenientes de vários trabalhos * | | | | | |
|-----------------------|---|------------|------------|------------|------------|-----------|
| % de pólen específico | 95/99; | >70; | 98±1; | >90 | | |
| pH | 4,2/4,8; | 4,4/4,8; | 3,7/4,8; | 4,8/ 6,8; | 4,6/5,4; | 4,6/6,5 |
| Acidez livre (meq/kg) | 23,0/32,0; | 31,5/56,5; | 24,4/72,8; | 9,3/11,3; | 10,3/24,8; | 13,4±3,3 |
| Frutose (g/100 g) | 35,0/44,7; | 34,7/41,2; | 35,1/43,5; | 37,1/47,4; | 41,9±1,3; | 37,6/46,0 |
| Glucose (g/100 g) | 25,2/30,8; | 25,6/30,5; | 28,8/30,6; | 21,0/30,0; | 22,1/29,3; | 23,6/32,4 |
| Sacarose (g/100 g) | 0/6,2; | 0,1/3,8; | 0/0,20; | 0,50±0,65; | 0,0/1,3; | 0,47/1,17 |
| Água (g/100 g) | 17,1/19,3; | 18,0/19,0; | 15,6/18,7; | 16,8/22,0; | 15,6/17,8; | 16,6/19,1 |

*) >mínimo; mínimo/máximo; média±desvio padrão

1.1.5. O mel em Portugal

A produção de mel é uma atividade agrícola com um peso económico razoável na economia nacional. De salientar que o sector apícola não deve ser avaliado somente pelo valor dos produtos diretos da atividade, como, por exemplo, o mel, a cera, o pólen, a própolis, a geleia real e as abelhas. Indiretamente, a ação da abelha na polinização e fertilização das plantas beneficia indiretamente outras produções agrícolas que superam este sector em valor [18].

Segundo a Federação Nacional dos Apicultores de Portugal (FNAP) e os dados apresentados no Programa Apícola Nacional para 2014-2016, os dados referentes à evolução desta atividade entre os anos 2010-2013 indicam que existem cerca de 17 mil apicultores registados, dos quais 673 são apicultores profissionais que detêm mais de 41% das 567 mil colmeias que estão contabilizadas no país. A produção de mel, devido à irregularidade do clima, corresponderá a uma produção média de cerca de 8 mil toneladas, o que equivale a uma faturação de cerca de 25 milhões de euros [18]. Em geral, 30% da produção nacional é para venda direta ao consumidor, 25% à indústria, 25% a centros de embalagem/comércio e 20% a retalhistas.

A produção é baixa comparativamente com a produção mundial, que se estima ser superior a 1.200.000 toneladas por ano. A China é o maior produtor, seguida da Argentina, México, Estados Unidos e Canadá. Esta produção tem tido uma tendência crescente nos

últimos 20 anos, a par de um aumento do consumo, devido ao aumento geral do nível de vida e também a um interesse maior em produtos naturais e saudáveis. Portugal, tal como, o resto da União Europeia, é deficitário em mel, sendo necessário incrementar a sua produção para não estar dependente da importação.

1.1.6. Controlo de qualidade

A preocupação no controlo de qualidade do mel é garantir que é autêntico no que diz respeito aos requisitos legais, sendo proibido acrescentar-lhe ou remover qualquer ingrediente [3]. O mel comercializado deve ser analisado de forma rotineira de modo a garantir a autenticidade do produto. A autenticidade deve ser avaliada ao nível da produção, de forma a prevenir a adulteração por adição de outros ingredientes alimentares (por exemplo, xaropes de sacarose ou melaços), bem como por tratamentos que levem à alteração do produto (fermentação, aquecimento e filtração); e, também, indicações erróneas quanto à origem geográfica e botânica [11].

O controlo de qualidade completo das amostras de mel inclui análises organoléticas e físico-químicas, bem como a análise de pólen.

A caracterização físico-química é efetuada pela análise dos seguintes parâmetros (procedimentos analíticos estabelecidos pela Comissão Internacional do Mel (IHC) [14]: humidade, condutividade elétrica, acidez livre, níveis de hidroximetilfurfural, atividade distásica, quantidade aparente de açúcares redutores (glucose e frutose), quantidade aparente de sacarose e matérias insolúveis em água.

A análise de pólen é importante nos estudos de autenticidade do mel ao nível da origem geográfica e botânica, mas requer técnicos muito experientes e especializados, o que dificulta a análise rotineira das amostras. Situação semelhante ocorre com as análises organoléticas, com a agravante de os técnicos sensoriais estarem limitados à análise de um pequeno número de amostras por dia.

Estas duas situações mostram a necessidade de encontrar metodologias analíticas que permitam auxiliar estes técnicos, identificando com nível elevado de confiança as amostras de mel monofloral. Foi neste âmbito que se decidiu investigar as capacidades analíticas de uma língua eletrónica potenciométrica com o objetivo de avaliar se esta pode ser uma ferramenta analítica auxiliar para classificar amostras de mel monofloral, deixando a atividade dos técnicos da análise organoléptica e polínológica para os meios de classificação duvidosa.

1.2. Língua eletrónica

Segundo a União Internacional da Química Pura e Aplicada (IUPAC) [19], uma língua eletrónica (LE) é um sistema de multi-sensores não específicos e de baixa seletividade, ou seja, constituído por sensores químicos com resposta reprodutível, de elevada estabilidade e sensibilidade cruzada a diferentes espécies em solução e que, por isso, fornecem perfis de sinais que traduzem informação global sobre as amostras. No processamento dos perfis de sinais usam-se métodos estatísticos multivariados para atingir os objetivos da análise, que podem ser o reconhecimento da composição qualitativa (classificação e discriminação de amostras), a análise quantitativa de multi-espécies químicas e a avaliação do sabor das amostras.

As LEs são sistemas com merecido interesse pois permitem adquirir informação global sobre amostras complexas que não é possível com outras metodologias analíticas, principalmente mimetizar a língua humana na apreciação de sabores. Esta informação varia de acordo com as características da LE usada.

1.2.1. Sensores químicos

O tipo de sensores usados nas LEs depende do princípio de deteção usado que, na área das técnicas electroanalíticas, pode ser a potenciometria [20-24], a voltametria [25-27], a amperometria [28] e a impedância electroquímica [29].

O sistema analítico usado neste trabalho é o potenciométrico, que tem diversas vantagens: é o mais simples, pode ser construído no laboratório (como neste trabalho), tem grande durabilidade, permite uma vasta seleção de reagentes na preparação dos sensores e facilidade na sua aplicação, sendo de utilização fácil, por exemplo, basta uma simples lavagem entre amostras para se proceder à análise seguinte.

Nos sistemas potenciométricos, vários tipos de sensores têm sido usados, sendo os multi-sensores constituídos só com um mesmo tipo de sensores ou com misturas de vários tipos, como por exemplo: sensores de vidros de calcogenetos [24,30], de vidro sensível a pH [24,30], membranas cristalinas [24], metais [24,31], membranas poliméricas de sensibilidade cruzada [21,22,24] e membranas seletivas a iões [23,24]. Alguns dos tipos de sensores referidos são seletivos, sendo usados nas LEs por permitirem uma diferenciação de amostras por meio de parâmetros químicos específicos. Mas, como já foi referido, o critério de seleção de um sensor para um sistema de multi-sensores que se pretende usar como LE, é ter baixa seletividade e alta sensibilidade cruzada. Esta característica consiste na capacidade de um sensor responder a um número diferente de espécies em solução e, num sistema de multi-sensores, implica que diferentes sensores vão produzir sinais diferentes para a mesma matriz, que constituem um perfil de sinais tipo “espectro” da amostra. O perfil

de sinais corresponde a uma impressão digital de multicomponentes da amostra, permitindo o uso de ferramentas multivariadas de estatística para análise qualitativa e quantitativa. As características de uma LE dependem quer dos materiais usados (composição) na preparação dos sensores, quer dos sensores usados na preparação do multi-sistema, sendo definidas pelos mesmos parâmetros que para um sensor discreto (seletividade, limite de deteção, limite de quantificação, etc). Na análise qualitativa, os trabalhos incidem no reconhecimento ou identificação de soluções.

O número de sensores num sistema de multi-sensores depende do trabalho que se pretende efetuar, tendo-se usado neste trabalho 40 sensores (20 sensores diferentes na sua composição, em duplicado) de sensibilidade cruzada. Os sistemas com um elevado número de sensores permitem um maior conjunto de aplicações porque, quer na classificação de amostras, quer na análise quantitativa de componentes em amostras, usando técnicas de seleção de variáveis acoplados às metodologias de análise estatística multivariada, possibilitam, selecionar os sensores mais adequados para o trabalho.

1.2.2. Processamento de sinais

O processamento dos sinais pode ser efetuado usando métodos estatísticos multivariados não supervisionados ou supervisionados.

Nos métodos não supervisionados estuda-se a variabilidade natural existente nas amostras analisadas com o objetivo de procurar algum padrão ou estrutura entre as amostras. Consistem em métodos exploratórios, sem qualquer controlo do investigador, que permitem identificar possíveis amostras estranhas e agrupamentos. A metodologia mais usada é a análise de componentes principais (ACP) [21,32]. Quando se pretende classificar ou quantificar compostos em amostras é necessário fornecer informação inicial sobre as amostras para que seja possível estabelecer modelos de classificação e de calibração, respectivamente, através da sua relação com os sinais da LE (métodos supervisionados). Dentro desta categoria, as metodologias mais usadas para classificação são a análise discriminante linear (ADL), o método dos mínimos quadrados parciais (MQP-AD) e as redes neurais artificiais (RNA) [32]. Para a análise quantitativa, os mais aplicados são a regressão linear múltipla (RLM) e o método dos mínimos quadrados parciais (MQP) [32,33].

Este trabalho incidiu na classificação de amostras através da ADL com o objetivo de, estabelecendo modelos matemáticos com os perfis de sinais da LE destas amostras, definir o tipo de mel monofloral. A ADL é uma técnica multivariada supervisionada de classificação que visa a obtenção de um modelo matemático resultante de combinações lineares das variáveis independentes (sinais dos sensores) que melhor separe os diferentes grupos (variável dependente) em que se pretendem classificar as amostras. A aplicação desta

metodologia implicou a execução de 2 passos: primeiro, o estudo de seleção das variáveis independentes (sensores) para obter o melhor modelo; depois, o estudo da robustez do modelo selecionado associado à classificação das amostras. Estes passos têm como objetivo definir as características que distinguem as amostras de um grupo de outro, de modo que, se possa prever a que grupo pertence uma nova amostra. Ao nível dos pressupostos de aplicação desta técnica, é importante verificar se há presença de *outliers*, normalidade nas variáveis, homogeneidade das matrizes de variância-covariância nos grupos e diferenças significativas entre os grupos. Mas, globalmente, a validação do modelo ADL é efetuada testando a sua capacidade de previsão pois, mesmo que um ou dois dos pressupostos não sejam observados, se for robusto, um modelo pode ser usado na previsão do grupo onde se inserem novas amostras [34]. O estudo da previsão, quando o número de amostras de cada grupo de classificação não é muito elevado, baseia-se principalmente nos resultados de classificação correta obtidos com a validação cruzada usando a técnica “deixar um de fora” (LOO, *leave-one-out*). Esta permite avaliar o desempenho do modelo na previsão ao remover uma amostra do estudo (contém n amostras) e prever a sua classificação com o modelo ADL obtido com as restantes ($n-1$). Este processo é repetido n vezes, de forma a obterem-se os erros de classificação de todas as amostras, sendo o erro global a soma dos erros de cada teste dividido por n .

Na seleção de variáveis (redução), os métodos habitualmente usados são os de seleção sequencial para a frente (*sequential forward selection method*), de seleção sequencial por eliminação posterior (*sequential backward elimination method*) e seleção por etapas (*stepwise method*) que permitem definir o modelo considerando a entrada e/ou remoção de uma variável de cada vez. Na técnica de seleção sequencial para a frente adiciona-se uma variável de cada vez (a primeira é a que tem maior correlação com a variável dependente) de forma a aumentar a validade da discriminação entre os grupos. Esta metodologia ignora as inter-relações entre as variáveis que ainda não foram selecionadas e dificilmente remove variáveis já adicionadas, o que pode causar grande redundância nas adições posteriores. Na técnica de seleção sequencial por eliminação posterior começa-se por considerar todo o conjunto de variáveis e, posteriormente, excluir variáveis, uma a uma, que contribuem menos para o poder de discriminação da função, mostrando as mesmas desvantagens do método acima referido. O método de seleção por etapas é uma combinação de ambas as técnicas, de seleção sequencial para a frente e seleção sequencial por eliminação posterior. Estas metodologias podem ser importantes quando se espera ordenar as variáveis selecionadas em termos da sua importância na diferenciação dos grupos mas, verifica-se que estes procedimentos nem sempre escolhem o melhor conjunto de variáveis independentes, principalmente quando o número destas é

elevado e mostram multicolinearidade (situação encontrada nos sensores de sensibilidade cruzada da LE). Estes métodos de seleção são chamados de métodos heurísticos (métodos de seleção de um subconjunto de variáveis) incluídos na classe chamada *greedy*, porque são métodos que fazem uma escolha no momento mas, mais tarde, não verificam se foi uma boa opção [35].

Por este motivo, deve-se verificar qual o melhor modelo de discriminação testando vários subconjuntos possíveis das variáveis, o que permite ter modelos com menor número de variáveis e eliminar variáveis redundantes ou que contribuem da mesma forma para as diferenças entre grupos, aumentando a exatidão de previsão ao eliminar variáveis irrelevantes. O modelo será mais simples, mais fácil de interpretar e, por se conhecerem as variáveis corretas, mais rápido nos cálculos, permitindo melhores resultados de previsão. Assim, neste trabalho, usaram-se algoritmos de seleção de variáveis chamados meta-heurísticos que identificam um subconjunto de variáveis originais que correspondem a um ótimo global para um dado critério de aproximação a uma determinada função, num grande espaço de busca de outros subconjuntos de variáveis.

Neste trabalho apresenta-se um estudo de comparação entre 3 algoritmos meta-heurísticos de busca local para a escolha de subconjuntos de variáveis baseados no pacote *Subselect* do programa R: algoritmo genético (*genetic algorithm*), algoritmo de arrefecimento simulado (*simulated annealing algorithm*) e algoritmo de melhoramento local restringido (*restricted local improvement algorithm*) [36,37].

O algoritmo **genético** (GEN) consiste em lidar com problemas de optimização simulando a evolução de uma população para a qual se deve definir a quantidade de indivíduos, a formação cromossómica do indivíduo e as probabilidades de aplicação dos operadores genéticos, tais como, cruzamento ou mutação. Nos indivíduos selecionados atuam os operadores genéticos (cruzamento e mutação) e os novos indivíduos substituem os anteriores, terminando, assim, uma geração. O algoritmo prossegue ciclicamente a partir dessa nova população e só termina quando algum critério de convergência é alcançado.

Mais detalhadamente, definindo-se um subconjunto de k variáveis para otimizar o problema, uma população de n subconjuntos de k variáveis é selecionado aleatoriamente a partir do conjunto total das variáveis independentes (p). Em cada iteração, $n/2$ casais são formados a partir da população e cada um gera um filho (um novo subconjunto de k variáveis), que herda as propriedades de seus pais (corresponde a todas as variáveis comuns a ambos os pais e uma seleção aleatória de variáveis na diferença simétrica da composição genética de seus pais). Cada descendência pode, opcionalmente, ser submetida a uma mutação na forma de um algoritmo de melhoria local (probabilidade especificada a definir no modelo). Os pais e filhos são selecionados de acordo com o seu

valor critério de qualidade, e o melhor desses k-subconjuntos irá tornar-se na próxima geração, que é usado como a população atual na iteração seguinte. O processo iterativo termina quando se atinge o número de gerações máximo definido inicialmente [36-38].

O algoritmo **arrefecimento simulado** (AS) é derivado de uma analogia com o processo de arrefecimento simulado da física de materiais, uma técnica que consiste em aquecer uma substância até à sua fusão seguida, de uma lenta diminuição da temperatura para que, a substância possa atingir a configuração mais regular possível (livre de defeitos) da sua malha cristalina (estado cristalino de energia mínimo).

O algoritmo estabelece a procura de um mínimo global para a optimização de um sistema com k variáveis considerando este tipo de comportamento termodinâmico. Em cada iteração, os valores de duas soluções (o corrente subconjunto de k variáveis e o novo subconjunto, também de k variáveis, que se pretende testar) são comparadas segundo um critério que mede a qualidade desses subconjuntos de variáveis. A nova solução é selecionada aleatoriamente na vizinhança da solução corrente e testada segundo as regras de arrefecimento simulado, tornando-se a solução corrente se o critério tiver valores superiores à inicial. O algoritmo prossegue a procura de novas soluções até atingir o número máximo de interações estabelecido no início do procedimento [36-38].

O algoritmo de **melhoramento local restringido** (MLR), um algoritmo de busca local alterado, é um procedimento iterativo que progressivamente melhora uma solução inicial por introdução de novas variáveis, por substituição uma a uma, procurando uma solução com valor maior do critério de qualidade (optimização por maximização). Se a nova solução não for encontrada, a solução vigente mantém-se e o algoritmo continua com o processo de encontrar novo máximo.

O algoritmo começa por estabelecer aleatoriamente um subconjunto inicial de k variáveis (número de variáveis definido *à priori* para a solução do problema), a partir das p variáveis independentes. As p-k variáveis que não entraram neste subconjunto são colocadas em “lista de espera”. O algoritmo vai substituindo uma variável no subconjunto de k corrente com uma variável em “lista de espera” e verifica quais das duas soluções origina os valores de critério mais elevados. Se a nova solução melhora o critério corrente, o subconjunto é atualizado. Nesta situação, a variável removida vai para a “lista de espera”, desde que seja a primeira vez (uma variável não pode entrar na “lista de espera” duas vezes). O algoritmo prossegue estas iterações até que a lista é esvaziada [36,37].

Estas técnicas de seleção de variáveis foram aplicadas aos dados analíticos obtidos com uma LE pela primeira vez neste trabalho, na esteira de uma pesquisa efetuada com vista a realizar o tratamento de dados, sendo também objetivo do trabalho averiguar a sua

eficácia para obter subconjuntos de sensores para os estudos de discriminação de méis monoflorais.

1.2.3. Exemplos de aplicação de LEs

As metodologias multivariadas de tratamento de dados obtidos por meio de LEs potenciométricas têm sido aplicadas em diferentes áreas científicas, principalmente na área alimentar, devido às evidentes vantagens associadas quer à construção dos sistemas de multi-sensores potenciométricos, quer à metodologia analítica. Por exemplo, os sistemas de multi-sensores são sistemas adaptáveis a uma grande variedade de situações, por haver variados tipos de sensores selecionáveis que são também compatíveis com a micro-fabricação, e que permitem a análise de matrizes líquidas complexas com um pré-processamento da amostra mínimo e por meio de um processo de medição simples. Estas LEs permitem unidades portáteis, deteção rápida, análise em tempo útil e economia na análise, também associada ao baixo custo do equipamento, que permite aplicar química analítica verde.

As LEs potenciométricas têm sido usadas principalmente na análise qualitativa de produtos na área alimentar. Estes sistemas têm sido usados, por exemplo, na distinção de bebidas [33,39] e alimentos [40,41], na classificação de vinho [42], na identificação de adulterações no leite [43] e na quantificação de sabores e alguns compostos [39,44].

Presentemente, o mel tem sido objeto de trabalho para LEs quer com base no princípio potenciométrico, quer no voltamétrico.

Quanto a estas últimas, foi usada uma LE voltamétrica, baseada na voltametria por impulsos, usando o eletrodo de referência Ag/AgCl, um contra-elektrodo de platina e seis metais como eletrodos de trabalho (ouro, prata, platina, paládio, tungsténio e titânio), num trabalho que possibilitou discriminar sete tipos de méis de diferentes origens florais (*Acacia*, *Astragali*, *Buckwheat*, *Coptis*, *Data*, *Motherwort* e *Vitex*) provenientes da mesma origem geográfica, a China [25]. Outro trabalho com LE voltamétrica, baseada num sistema de três eletrodos com eletrodo de trabalho de platina, eletrodo de referência Ag/AgCl e contra-elektrodo de platina, usando voltametria cíclica, permitiu identificar 4 tipos de mel de diferentes origens botânicas (*Eucalyptus*, *Til*, *Leechi* e *Khalisa*), da Índia, com elevada exatidão [26].

Com LEs potenciométricas, salientam-se quatro trabalhos: o primeiro, usando um sistema de multi-sensores contendo 20 sensores preparados com membranas poliméricas lipídicas e um eletrodo de referência Ag/AgCl, permitiu discriminar de forma satisfatória 3 tipos de méis monoflorais (méis de *Erica*, *Echium* e *Lavandula*) Portugueses [21]; no segundo, aplicou-se uma LE comercial com 7 sensores e um eletrodo de referência

Ag/AgCl, para classificar 3 tipos de méis (*Acacia*, *Castanea* e melada) da Croácia, com elevada exatidão (100% de classificações corretas) [45]; no terceiro, usando o mesmo sistema comercial, 8 tipos de mel chineses (*Acacia*, *Astragali*, *Data*, *Coptis*, *Vitex*, *Motherwort*, *Radix Changli* e *Buckwheat*) foram classificadas de forma eficiente (>90% de classificações corretas) [46]; por último, um sistema constituído por 3 elétrodos metálicos puros e por 4 elétrodos de compostos metálicos foi usado para discriminar méis de *Citrus*, *Rosemary*, multifloral e melada, de forma eficiente (>93% de classificações corretas) [47].

1.3. Génese do trabalho

Neste trabalho utilizou-se uma LE potenciométrica contendo membranas lipo/poliméricas, de sensibilidade cruzada, com o objetivo de desenvolver alternativas de análise para a classificação de mel que é efetuada através da identificação e quantificação de pólenes presentes na amostra, uma técnica biológica demorada e que requer trabalho especializado. Este estudo é a progressão de um trabalho anterior [21] desenvolvido pela mesma equipa de investigação, em que se usou um sistema de multi-sensores construído em acrílico e com membranas poliméricas de composição diferente para discriminar 3 tipos de méis monoflorais (*Echium*, *Erica* e *Lavandula*). Conforme já referido, o objetivo global do presente estudo de tese de mestrado foi investigar se uma nova LE construída permite classificar corretamente as amostras de mel nos seus grupos monoflorais. Os objetivos específicos foram já listados atrás.

2. Material e métodos

2.1. Amostragem

As 89 amostras de mel usadas neste trabalho foram fornecidas pela Federação Nacional de Apicultores de Portugal (FNAP). Estas amostras referem-se a méis que entraram em concursos de mel no âmbito dos Fórum Nacional de Apicultura/Feira Nacional do Mel organizados pela FNAP, em 3 anos consecutivos (2010-2012), e que se referem aos anos apícolas 2009 a 2011 (respectivamente 29, 37 e 23 amostras). As amostras são provenientes de todas as regiões apícolas do território continental de Portugal.

As amostras foram analisadas usando vários procedimentos, a seguir descritos, de forma a obter-se uma caracterização sensorial (ao nível da cor) da composição polínica e química (ao nível dos açúcares), para ser relacionada com a informação obtida da análise do mel com a LE.

2.2. Caracterização da cor do mel

A cor foi caracterizada numa escala de mm PFund, a partir da medição da absorvância de uma solução aquosa de cada amostra num espectrofotómetro UV-Vis (Jenway, modelo Genova) a um comprimento de onda de 635 nm [48]. A solução aquosa foi preparada por diluição da amostra na proporção de 1:2 (5,0 g de mel em 10,0 mL de água desionizada). No caso de haver evidência de presença de cristais de açúcares, a amostra foi aquecida a 50°C previamente à dissolução. A solução foi deixada repousar na cuvete de quartzo durante 10 minutos antes da medição. Como branco usou-se água desionizada.

A cor do mel foi expressa em unidades da escala de mm Pfund aplicando a seguinte expressão [48]:

$$mm\ PFund = -38,70 + 371,39 \times Absorvância \quad (Eq. 1)$$

Para verificar visualmente os resultados usou-se também uma escala de cores padrão do mel impressa, associada a intervalos de valores de cor na escala mm PFund, apresentada na Figura 1.

Na classificação da cor usou-se a escala definida pelo Departamento de Agricultura dos Estados Unidos (USDA) [49] que considera 7 grupos de cor para o mel. Na Tabela 2 mostra-se a correspondência de cada cor com os intervalos de cor da escala mm PFund. Nos méis portugueses, a cor branco água não é usual e o mel de cor âmbar escuro é referido como mel escuro.

Tabela 2 – Cores padrão do mel designadas pela USDA e respetivos intervalos de cor da escala mm PFund.

| Cores padrão do mel | Intervalos de cor da escala mm PFund |
|---------------------|--------------------------------------|
| Branco água | ≤ 8 |
| Extra branco | >8 e ≤ 17 |
| Branco | >17 e ≤ 34 |
| Âmbar extra claro | >34 e ≤ 50 |
| Âmbar claro | >50 e ≤ 85 |
| Âmbar | >85 e ≤ 114 |
| Âmbar escuro * | >114 |

*) Geralmente classificado com mel escuro

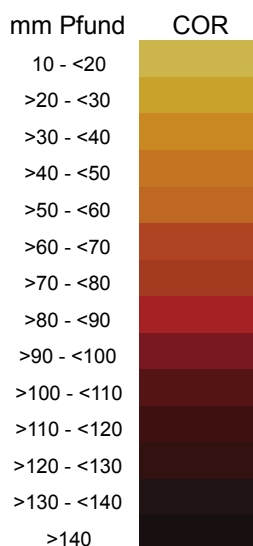


Figura 1 – Escala impressa de cores padrão do mel e respetivos intervalos de valores mm PFund.

2.3. Análise polínica

Na análise polínica (quantitativa), as amostras foram preparadas usando o método de Louveaux [50]. As medições polínicas foram efetuadas a partir de 10,0 g de amostra, dissolvidas em água destilada e o sedimento foi concentrado através de centrifugações a 1500 rpm durante 30 minutos. Ao sedimento adicionaram-se 10,0 mL de uma mistura de anidrido acético e ácido sulfúrico (9:1). Após incubação em banho-maria (100 °C durante 3 min) com agitação manual e centrifugação, a solução foi decantada. Posteriormente, adicionaram-se cerca de 12,0 mL de ácido acético ao sedimento e, após agitação, efectuou-se uma nova centrifugação e decantação. O sedimento foi lavado e ressuspensão em 12,0 mL de água destilada, sendo novamente centrifugado e decantado. A lavagem final foi

efetuada com 12,0 mL de hidróxido de potássio 7%, repetindo-se os passos de agitação, centrifugação e decantação. Em seguida, os grãos de pólen foram corados com uma solução de fucsina básica (Merck) e misturada com glicerina (Absolve).

A observação microscópica dos pólenes foi efetuada com um microscópio (Leitz Messtechnik GmbH, Wetzlar, Germany), com objetivas 400x e 1000x (esta última amplificação foi usada quando foi necessário maior pormenor da estrutura do pólen para uma clara identificação).

Os pólenes presentes em cada amostra de mel foram identificados por uma investigadora com bastante experiência em análise polínica (foi contado um mínimo de 1000 pólenes por cada amostra, tendo-se repetido a análise até a obtenção de valores concordantes), usando como base de comparação uma coleção de referência da Escola Superior Agrária - Instituto Politécnico de Bragança e manuais de morfologia do pólen. Os grãos de pólen foram agrupados em tipos polínicos (conjunto de grãos de pólen com morfologia semelhante, normalmente originários de plantas do mesmo género).

2.4. Análise por HPLC

A cromatografia líquida de alto desempenho (HPLC) foi usada neste trabalho para identificar e determinar a concentração dos açúcares principais presentes no mel: glucose e frutose.

2.4.1. Equipamento

O equipamento de cromatografia líquida, da marca Varian, era constituído por uma bomba Prostar 220, um injetor manual Rheodyne modelo 7725i com um loop de 20 µL, um forno Jones CHromatography modelo 7981 e uma coluna Supelcogel C-610H (30 cm de comprimento e 7,8 mm de diâmetro interno). Na análise e quantificação dos açúcares (glucose e frutose) usou-se o detetor índice de refração (IR) da Varian modelo RI-4. O fluxo do eluente, em modo isocrático, foi de 0,4 mL/min. O tempo de análise de cada solução foi de 30 minutos.

2.4.2. Preparação de soluções

Eluente. Na análise de açúcares, o eluente usado foi ácido fosfórico (Fisher Scientific) a 0,1%. Todos os eluentes preparados para o HPLC foram previamente filtrados com um sistema de filtração da Phenomenex, acoplado a uma bomba de vácuo da Laboport, usando filtros 0,2 µm de nylon da marca Whatman, e desgasificados em ultrassons (Elma Transsonic 460/H) durante aproximadamente 5 minutos.

Soluções padrão. Para definir os tempos de retenção de vários açúcares referentes à análise por HPLC, prepararam-se soluções independentes dos seguintes açúcares, de concentração $1,00 \times 10^{-3}$ mol/L: D(+)-trealose (Fluka), maltose (Panreac), D(+)-manose (Merck), alfa-D(+)-melibiose (Riedel-de Haen), D(+)-turanose (ACROS Organics), D(+)-glucose (Fluka), D(+)-melezitose (Sigma), D(-)-frutose (Fluka) e sacarose (Panreac).

Na calibração do equipamento usaram-se soluções padrão de calibração de mistura da sacarose, frutose e glucose. Estas soluções aquosas foram preparadas por medição de massas de cada composto diretamente para balões volumétricos, de forma a obter concentrações entre 0,500 e 5,00 g/L.

Para o estudo da precisão e exatidão prepararam-se duas soluções de controlo de qualidade usando o mesmo procedimento que para os compostos acima referidos, com níveis de concentração baixo e médio, quer para a glucose (valores aproximados de 0,74 g/L e 2,10 g/L, respectivamente), quer para a frutose (valores aproximados de 0,75 g/L e 3,2 g/L, respectivamente).

Amostras. Todas as amostras foram diluídas por medição de 0,30-0,50 g de mel para um balão volumétrico de 50,00 mL, aferido com água desionizada. As soluções foram filtradas com o auxílio de microfiltros descartáveis (nylon 0,2 μ m da Whatman) e seringas para dentro de vials, para serem posteriormente analisadas por HPLC.

No estudo da repetibilidade e precisão intermédia foram usadas três amostras de mel.

2.5. Análise com a língua eletrónica

2.5.1. Montagem do sistema de multi-sensores

O sistema de multi-sensores foi preparado numa placa de cloreto de polivinílico (PVC) através da impressão, nas duas faces, de um circuito usando o método *print-screen*, por aplicação de uma pasta epoxilica condutora de prata (EPO-TK E4110, Epoxy Technology, Inc.) de cura a baixas temperaturas, preparada por mistura de dois reagentes (pasta + endurecedor). Em cada face do sistema é possível aplicar 10 sensores químicos, ou seja, cada sistema permite monitorizar o sinal de 20 sensores.

A temperatura de cura usada foi de 40°C, permitindo a secagem da pasta em 8 horas. Após corte e limpeza do circuito impresso, foram vedados os pontos de colocação do sensor químico e de ligação à ficha RS-232, para posterior impermeabilização do sistema, usando como revestimento uma resina acrílica (PLASTIK 70, Kontakt Chemie) adicionada por pulverização. A fase final consistiu na colocação da ficha RS-232 de 25 pinos no sistema, envolvendo a sua fixação e ligação ao circuito impresso (Figura 2). Cada sistema

foi testado com um multivoltímetro para confirmar se todas as ligações permitiam a condução de sinal elétrico.

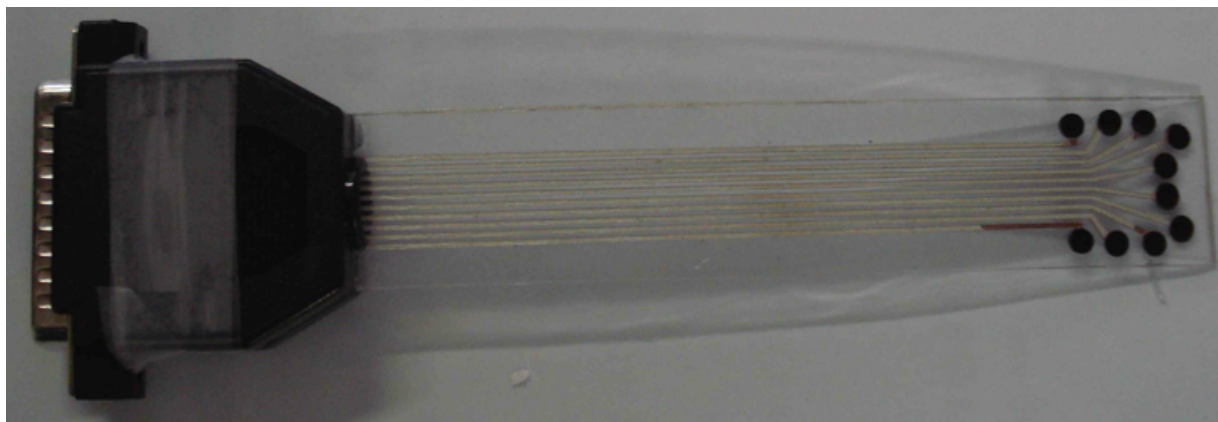


Figura 2 – Sistema de multi-sensores.

2.5.2. Sensores químicos

Os sensores químicos usados correspondem a diferentes membranas de sensibilidade cruzada, preparadas com diferentes combinações de vários compostos aditivos lipídicos (entre 2,8 e 3,2%) e plastificantes (entre 64,7 e 65,2%). O polímero usado foi o PVC (entre 31,9 e 32,3%) da Fluka. Usaram-se 4 substâncias aditivas: octadecilamina, álcool oleílico, cloreto de metiltriocetilamônio e ácido oleico (todos Fluka). As 5 substâncias plastificantes usadas foram: adipato de bis-(1-butilpentilo), sebacato de dibutilo, 2-nitrofenil octil éter, fosfato de 2-etil-hexilo e fosfonato de dioctilfenilo (todos Fluka).

Cada mistura foi preparada por medição de massas pré-estabelecidas de cada um dos três produtos (ver Tabela 3) e dissolvidos com o solvente tetrahydrofurano (Sigma), de forma a obter-se uma solução homogênea e viscosa.

A formação de cada membrana no sistema de multi-sensores foi efetuada usando a técnica da gota (adições de uma gota a intervalos de tempo de 3-5 minutos para evaporação completa do solvente) até à formação de uma membrana polimérica transparente e de aparência cristalina.

2.5.3. Equipamento

A língua eletrônica utilizada neste trabalho é um sistema analítico constituído por eléctrodos potenciométricos “tudo-estado-sólido”, o que lhe confere robustez na sua utilização. É um dispositivo constituído por dois sistemas de multi-sensores, iguais na sua constituição ao nível dos sensores químicos, e um eléctrodo de referência Ag/AgCl de dupla

junção. A língua eletrónica estava ligada a um datalogger (Agilent 34970). A aquisição do sinal de cada membrana foi efetuada através de um computador usando o software “Agilent BenchLink Data Logger” (Figura 3). As análises foram efetuadas em soluções sob agitação magnética (VELP Scientifica, modelo ARE Heating magnetic stirrer).

Tabela 3 – Composição das membranas poliméricas.

| Plastificante | Aditivo | Sensor nº | % Aditivo | % Plastificante | PVC |
|---------------------------------|--------------------------------|-----------|--------------|--------------------|-------|
| Adipato de bis-(1-butilpentilo) | Octadecilamina | S1 | 2,99 | 65,02 | 31,99 |
| | Álcool oleílico | S2 | 2,96 | 65,00 | 32,04 |
| | Cloreto de metiltrioctilamónio | S3 | 3,00 | 65,01 | 31,99 |
| | Ácido oleico | S4 | 3,00 | 65,04 | 31,96 |
| Sebacato de dibutilo | Octadecilamina | S5 | 3,02 | 65,11 | 31,87 |
| | Álcool oleílico | S6 | 2,96 | 65,02 | 32,02 |
| | Cloreto de metiltrioctilamónio | S7 | 3,00 | 64,91 | 32,10 |
| | Ácido oleico | S8 | 2,99 | 64,93 | 32,09 |
| 2-nitrofenil octil éter | Octadecilamina | S9 | 2,98 | 64,68 | 32,34 |
| | Álcool oleílico | S10 | 3,00 | 65,10 | 31,90 |
| | Cloreto de metiltrioctilamónio | S11 | 2,99 | 65,00 | 32,00 |
| | Ácido oleico | S12 | 3,01 | 65,04 | 31,95 |
| Fosfato de 2-etil-hexilo | Octadecilamina | S13 | 2,96 | 64,99 | 32,05 |
| | Álcool oleílico | S14 | 2,97 | 65,04 | 32,00 |
| | Cloreto de metiltrioctilamónio | S15 | 3,00 | 64,99 | 32,01 |
| | Ácido oleico | S16 | 2,98 | 65,04 | 31,98 |
| Fosfonato de dioctilfenilo | Octadecilamina | S17 | 3,12 | 64,85 | 32,03 |
| | Álcool oleílico | S18 | 2,95 | 65,04 | 32,01 |
| | Cloreto de metiltrioctilamónio | S19 | 2,99 | 64,69 | 32,32 |
| | Ácido oleico | S20 | 3,01 | 65,11 | 31,88 |

2.5.4. Análise com a língua eletrónica

Para analisar as amostras com a LE prepararam-se soluções aquosas de mel por diluição de 10,00 g de mel com 50,00 g de água desionizada.

As análises implicam só um passo prévio de lavagem cuidadosa do sistema analítico com água desionizada e enxaguamento leve com papel absorvente, antes de ser mergulhado na nova solução a analisar. O sistema permite a monitorização contínua dos sensores, pelo que foi possível verificar visualmente a variação do sinal em função do tempo de cada uma das 40 membranas poliméricas usadas. Cada análise implicou um período de

estabilização do sinal de 7 minutos, com agitação magnética, findo o qual se removia a LE da solução medida para a preparar para nova análise.

Todas as experiências foram efetuadas à temperatura ambiente.

Os perfis de sinais de potencial de cada membrana polimérica foram transferidos para o programa Excel para posterior tratamento de dados.

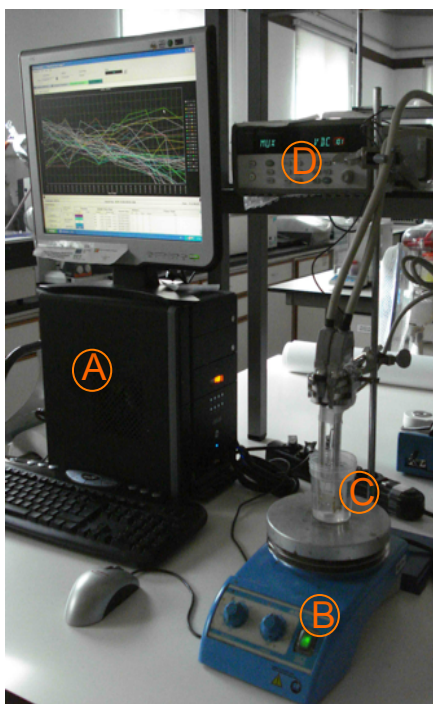


Figura 3 – Montagem para a análise com a língua eletrônica.

A - Computador; B - Agitador magnético; C - Sistemas de multi-sensores + elétrodo de referência inserido numa amostra; D - Datalogger

2.6. Análise de dados

Os dados multivariados obtidos experimentalmente neste trabalho foram tratados com o programa de estatística R (versão 2.15.1). O estudo incidiu em estabelecer um procedimento de tratamento de dados de forma a permitir classificar corretamente, usando o método estatístico multivariado ADL (técnica supervisionada de classificação), o grupo de mel monofloral de um elevado número de amostras com intrínseca variabilidade. Considerando que os perfis de sinais obtidos com a LE têm uma elevada multicolinearidade entre sensores, os modelos ADLs obtidos são avaliados atendendo à sua robustez ao nível da previsão, o que permite a comparação dos modelos, tendo em conta as classificações corretas das amostras considerando os dados originais e, principalmente, as obtidas usando a técnica de validação cruzada LOO (*leave-one-out*). Em simultâneo com este

procedimento, usaram-se três algoritmos meta-heurísticos de seleção de subconjuntos de variáveis por busca local (disponíveis no pacote *Subselect* do programa R) para efeitos de comparação e para estabelecer qual o que permitia obter os melhores modelos de classificação ADL: algoritmo genético (GEN), algoritmo de arrefecimento simulado (AS) e algoritmo de melhoramento local restringido (MLR). Nas condições definidas, cada algoritmo apresentou o melhor modelo para cada subconjunto a variar entre 2 e 20 sensores e após 10.000 tentativas de obter esse modelo. Este número foi estabelecido por se ter verificado que permitia chegar à mesma solução nos subconjuntos com mais de 15 sensores, ou seja, ao mesmo conjunto de sensores selecionados. O critério de qualidade usado para avaliar o ajuste entre a variável dependente (grupos de mel monofloral) usando os subconjuntos de sensores estabelecidos pelos algoritmos foi o critério *ccr12* (*coefficient and Roy's first root statistic*, ccr_1^2) [36,37]. A maximização deste índice é equivalente à maximização da primeira raiz de Roy, que corresponde à proporção entre a variância explicada e a não explicada para a primeira função discriminante (valor conceitualmente equivalente à razão F na análise de variância univariada).

Os 19 modelos obtidos com cada um dos três algoritmos de seleção usados foram, posteriormente, testados para inferir da sua capacidade de classificação das amostras de mel de acordo com os diferentes tipos de méis monoflorais (quer com dados originais quer com a metodologia validação cruzada LOO), selecionando-se o que permitia um maior número de classificações corretas para a validação cruzada LOO.

Um primeiro estudo envolveu a classificação de todas as amostras de mel monofloral (excluindo as amostras de mel multifloral), usando a ADL. Em seguida, efectuou-se um estudo semelhante sem incluir também os grupos de tipo de mel representadas com menos de 2 amostras para averiguar se permitia obter modelos de previsão mais robustos. Finalmente, definiu-se a metodologia a seguir, ao nível prático, para a discriminação de amostras de diferentes tipos de méis com pelo menos 3 ou mais amostras. Esta metodologia envolveu o uso da cor da amostra como parâmetro primário de seleção de amostras para simplificar o tratamento estatístico dos dados analíticos e permitir obter classificações corretas através da análise multivariada. A solução envolveu a divisão das amostras em três grupos de cores (detalhes apresentados no capítulo seguinte) e tratamento de dados de forma independente em cada grupo.

3. Resultados

3.1. Cor do mel

A cor não é um parâmetro de qualidade do mel, mas é um factor comercial importante para os produtores, embaladores e consumidores. Em Portugal, não é necessário especificar a cor no rótulo, mas os consumidores baseiam-se neste atributo visual na escolha do mel que compram. Como a cor é uma característica que advém principalmente da fonte floral do mel, decidiu-se analisar os dados da análise polínica e verificar se é possível obter uma relação entre os respetivos resultados e os da análise da cor.

As 89 amostras de mel foram analisadas por espectrofotometria UV-Vis e determinada a cor na escala Pfund (mm) através das absorvâncias obtidas (equação 1). Na Figura 4 apresenta-se a classificação de cada amostra de acordo com a escala de cores padrão do mel impressa e a definida pela USDA. Na Tabela 4 mostra-se o número de amostras e o respectivo valor de percentagem associado a cada grupo de cor padrão definida pela USDA.

Tabela 4 – Número de amostras com cores padrão do mel designadas pela USDA.

| Cores padrão do mel | Nº de amostras | % |
|-------------------------|----------------|------|
| Extra branco (EB) | 2 | 2,2 |
| Branco (B) | 20 | 22,5 |
| Âmbar extra claro (AEC) | 10 | 11,2 |
| Âmbar claro (AC) | 22 | 24,7 |
| Âmbar (A) | 10 | 11,2 |
| Escuro (E) | 25 | 28,1 |

Como se pode visualizar na Figura 4, as amostras mostram uma variação de cor desde o extra branco até ao escuro, com variações dentro do intervalo de 15,9 a 204,2 mm Pfund. Segundo a escala definida pela USDA, as amostras de cor branco, âmbar claro e escuro são as que predominam (>22% em cada categoria). Méis de cor âmbar extra claro e âmbar estão representados medianamente (11,2% para cada categoria). As amostras de cor extra branco são as menos representativas (2,2%) .

Na Figura 5 mostra-se o gráfico da distribuição da cor Pfund das amostras e a função que melhor se ajustou à frequência das amostras, a distribuição log-normal. Esta distribuição mostra que o logaritmo da frequência segue uma distribuição normal.

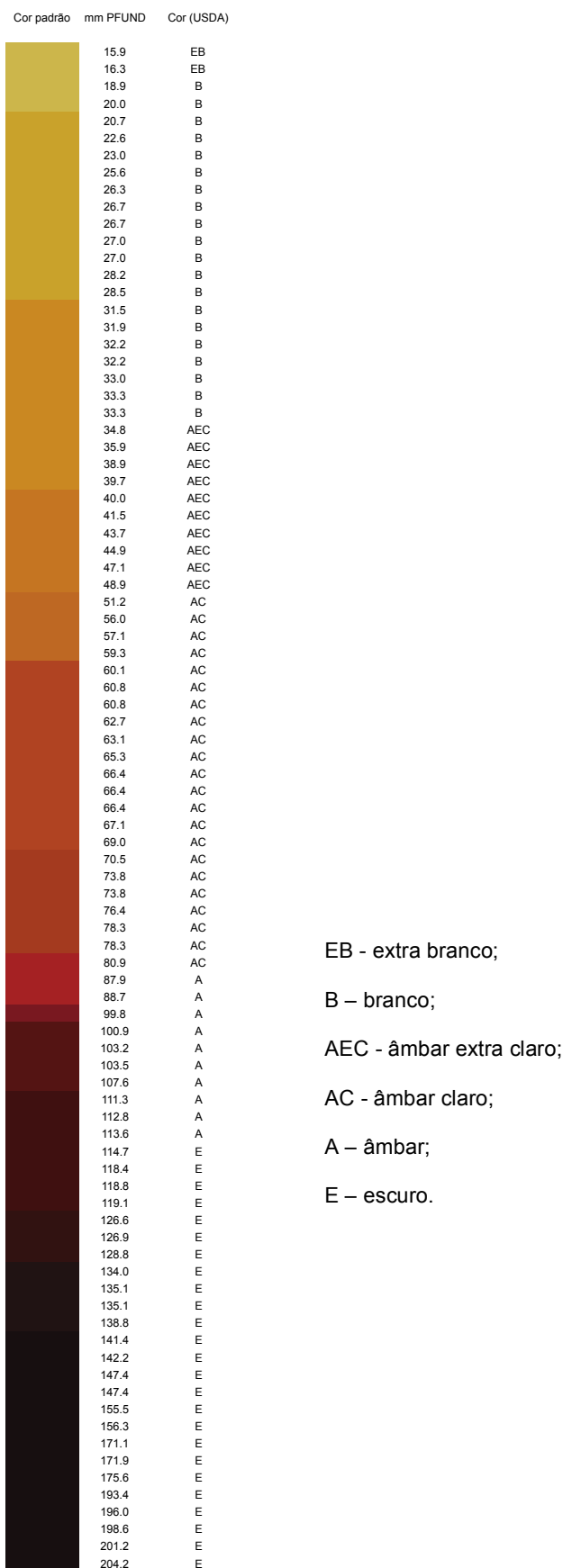


Figura 4 – Amostras de mel usadas neste trabalho de acordo com a escala de cores padrão do mel impressa e as definidas pela USDA.

A variabilidade da cor obtida é uma indicação de acentuada variação na composição química e, por isso, da fonte polínica das amostras. Esta variabilidade das amostras era pretendida com vista a testar as capacidades analíticas da LE nos estudos do mel.

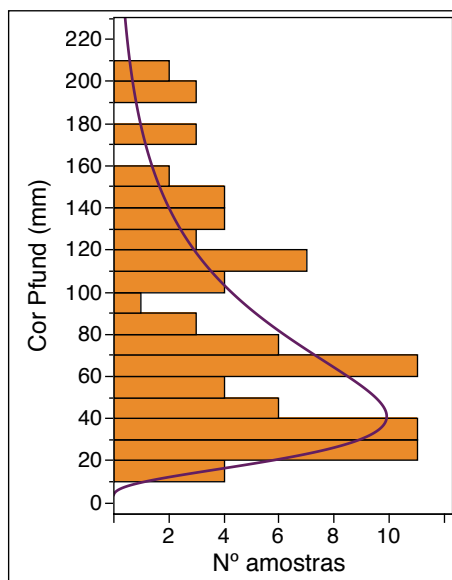


Figura 5 – Distribuição da cor Pfund (mm) e ajuste da função de distribuição log-normal.

3.2. Análise polínica

A análise polínica realizada permitiu obter o perfil polínico das amostras. Esta é uma análise quantitativa, uma vez que se faz a contagem dos grãos de pólen com morfologia semelhante, originários de plantas do mesmo género (por tipo de pólen). A proporção de cada tipo de pólen presente na amostra permite a sua classificação em termos de origem botânica, embora com limitações associadas à interpretação dos perfis polínicos. Em rigor, a classificação deveria ser complementada com uma análise sensorial.

Na Tabela 5 mostram-se os diferentes tipos de pólen visualizados nas amostras, a indicação da percentagem de méis com o pólen presente na sua constituição (ordem decrescente) e os valores mínimo e máximo da percentagem de cada pólen obtido da análise. Na Tabela 6 apresentam-se os nomes científicos do género da planta cujo tipo de pólen foi visualizado na análise polínica das amostras de mel e exemplos de plantas com nomes comuns, associadas ao género.

Na globalidade das amostras foi possível identificar 23 tipos de pólen. A ordem decrescente dos valores de percentagem de méis com um dado tipo de pólen mostra que os pólenes *Rubus sp.*, *Prunus sp.*, *Echium sp.*, *Lavandula sp.*, *Castanea sp.*, *Trifolium sp.*, *Erica*

sp. e *Eucalyptus sp.* estão presentes em mais de 40% das amostras. A seguir, os pólenes *Thymus sp.*, *Leontodeon sp.*, *Cytisus sp.*, *Acacia sp.* e *Pinus sp.* são os mais abundantes, estando presentes em 13 a 25% das amostras. Os restantes tipos de pólen estão presentes em menos de 8% das amostras.

Tabela 5 – Percentagem de méis com cada tipo de pólen presente e o valor mínimo e máximo da percentagem de cada pólen nas amostras de mel.

| Tipo de pólen | Abreviatura do tipo de pólen | % Pólen nas amostras | | % Méis com o pólen |
|-----------------------|------------------------------|----------------------|------|--------------------|
| | | Min | Máx | |
| <i>Rubus sp.</i> | Rub | 1,1 | 68,6 | 76,4 |
| <i>Prunus sp.</i> | Pru | 1,4 | 80,1 | 75,3 |
| <i>Echium sp.</i> | Ech | 1,4 | 79,3 | 73,0 |
| <i>Lavandula sp.</i> | Lav | 1,2 | 66,5 | 70,8 |
| <i>Castanea sp.</i> | Cas | 2,5 | 94,8 | 62,9 |
| <i>Trifolium sp.</i> | Tri | 1,4 | 81,7 | 50,6 |
| <i>Erica sp.</i> | Eri | 1,2 | 81,7 | 41,6 |
| <i>Eucalyptus sp.</i> | Euc | 1,2 | 82,0 | 40,4 |
| <i>Thymus sp.</i> | Thy | 1,3 | 17,8 | 24,7 |
| <i>Leontodeon sp.</i> | Leo | 1,1 | 20,1 | 24,7 |
| <i>Cytisus sp.</i> | Cyt | 1,2 | 7,1 | 19,1 |
| <i>Acacia sp.</i> | Aca | 0,6 | 21,8 | 14,6 |
| <i>Pinus sp.</i> | Pin | 1,6 | 6,1 | 13,5 |
| <i>Sandix sp.</i> | San | 1,5 | 6,6 | 7,9 |
| <i>Foeniculum sp.</i> | Foe | 1,5 | 24,3 | 7,9 |
| <i>Helianthus sp.</i> | Hel | 1,4 | 6,9 | 4,5 |
| <i>Mentha sp.</i> | Men | 1,5 | 3,0 | 3,4 |
| <i>Mimosaceae sp.</i> | Mim | 2,6 | 5,4 | 2,2 |
| <i>Crepis sp.</i> | Cre | 1,8 | 3,3 | 2,2 |
| <i>Tilia sp.</i> | Til | 4,7 | 11,2 | 2,2 |
| <i>Genista sp.</i> | Gen | 1,2 | 11,4 | 2,2 |
| <i>Medicago sp.</i> | Med | 19,9 | 19,9 | 1,1 |
| <i>Persea sp.</i> | Per | 6,9 | 6,9 | 1,1 |

Os diferentes tipos de pólenes encontrados são comuns nos méis Portugueses [12]. Os valores percentuais mínimos e máximos de cada tipo de grão de pólen indicam que o espectro polínico de cada amostra é variável, tal como o número de tipos de pólen presentes em cada amostra. Os perfis polínicos mostram que estas apresentam geralmente

4 a 9 tipos de pólen na sua constituição. Só quatro amostras apresentaram 3 tipos de pólen e apenas uma amostra continha dois tipos.

Tabela 6 – Exemplo de planta associada a cada género identificado através da análise polínica.

| Género da planta | Exemplo de planta (nomes triviais) |
|-----------------------|------------------------------------|
| <i>Acacia sp.</i> | Acácia |
| <i>Castanea sp.</i> | Castanheiro |
| <i>Crepis sp.</i> | Almeirão-branco |
| <i>Cytisus sp.</i> | Giesta |
| <i>Echium sp.</i> | Soagem |
| <i>Erica sp.</i> | Urze |
| <i>Eucalyptus sp.</i> | Eucalipto |
| <i>Foeniculum sp.</i> | Funcho |
| <i>Genista sp.</i> | Tojo |
| <i>Helianthus sp.</i> | Girassol |
| <i>Lavandula sp.</i> | Rosmaninho |
| <i>Leontodon sp.</i> | Dente-de-leão |
| <i>Medicago sp.</i> | Alfafa |
| <i>Mentha sp.</i> | Hortelã |
| <i>Mimosaceae sp.</i> | Mimosa |
| <i>Persea sp.</i> | Loureiro-real |
| <i>Pinus sp.</i> | Pinheiro |
| <i>Prunus sp.</i> | Cerejeira |
| <i>Rubus sp.</i> | Amora silvestre |
| <i>Thymus sp.</i> | Tomilho |
| <i>Tilia sp.</i> | Tília |
| <i>Trifolium sp.</i> | Trevo |
| <i>Sandix sp.</i> | Agulha de pastor |

Verifica-se também que os pólenes *Cytisus sp.*, *Pinus sp.*, *Sandix sp.*, *Helianthus sp.*, *Mentha sp.*, *Mimosaceae sp.*, *Crepis sp.* e *Persea sp.* estão presentes no mel com valores percentuais máximos inferiores a 7.2%. Esta baixa percentagem pode indicar que a contribuição destes tipos de pólen para a cor e sabor das amostras será pouco significativa.

A classificação das amostras em mel monofloral/multifloral foi efetuada usando como critério a predominância dos pólenes. Na Tabela 7 apresentam-se os resultados desta classificação e os três pólenes mais predominantes ($\geq 10\%$ de presença) identificados em cada amostra.

Tabela 7 – Classificação e os 3 pólenes mais predominantes das amostras.

| Classificação | 1° pólen - 2° pólen - 3° pólen * | |
|--------------------------|----------------------------------|--------------------------------|
| <i>Castanea</i> (≥90%) | Cas (90) | Cas (91) |
| | Cas (92) | Cas (95) |
| <i>Echium</i> (≥45%) | Ech (49) – Rub (16) | Ech (49) – Rub (19) – Cas (15) |
| | Ech (53) – Pru (22) – Foe (16) | Ech (53) – Rub (12) |
| | Ech (53) – Cas (16) | Ech (53) – Rub (28) |
| | Ech (54) – Pru (24) – Lav (10) | Ech (54) – Pru (16) |
| | Ech (55) – Cas (27) | Ech (58) |
| | Ech (58) – Cas (26) – Tri (13) | Ech (63) – Pru (26) |
| | Ech (70) – Pru (13) | Ech (70) – Lav (10) |
| | Ech (71) – Pru (15) | Ech (73) |
| | Ech (79) – Lav (12) | |
| <i>Erica</i> (≥45%) | Eri (50) – Cas (28) | Eri (53) – Cas (15) – Rub (13) |
| | Eri (54) – Cas (19) | Eri (54) – Cas (26) |
| | Eri (56) – Tri (23) | Eri (61) – Tri (17) – Pru (12) |
| | Eri (63) – Rub (19) – Cas (15) | Eri (82) – Cas (12) |
| <i>Eucaliptus</i> (≥70%) | Euc (72) – Rub (12) | Euc (82) |
| <i>Lavandula</i> (≥15%) | Cas (38) – Ech (29) – Lav (17) | Ech (34) – Rub (32) – Lav (19) |
| | Ech (36) – Lav (31) – Tri (21) | Ech (38) – Rub (25) – Lav (19) |
| | Ech (39) – Lav (21) – Rub (20) | Ech (39) – Rub (26) – Lav (21) |
| | Ech (60) – Lav (17) – Tri (14) | Lav (31) – Ech (29) – Pru (22) |
| | Lav (31) – Pru (29) – Ech (20) | Lav (32) – Cas (19) – Rub (13) |
| | Lav (33) – Rub (20) – Thy (18) | Lav (38) – Ech (30) – Rub (18) |
| | Lav (39) – Ech (39) | Lav (45) – Ech (31) |
| | Lav (45) – Ech (25) – Rub (20) | Lav (46) – Pru (40) |
| | Lav (47) – Cas (17) | Lav (49) – Rub (20) – Ech (10) |
| | Lav (56) – Leo (20) – Rub (15) | Lav (67) – Eri (12) – Rub (12) |
| | Pru (32) – Lav (27) – Ech (26) | Rub (39) – Lav (19) – Tri (11) |
| | | |
| <i>Prunus</i> (≥45%) | Pru (49) – Cas (30) – Ech (18) | Pru (50) – Cas (38) |
| | Pru (64) – Ech (15) – Aca (13) | Pru (80) |
| <i>Rubus</i> (≥45%) | Rub (47) – Euc (21) – Pru (19) | Rub (49) – Cas (17) – Gen (11) |
| | Rub (49) – Tri (17) – Cas (14) | Rub (49) – Cas (22) |
| | Rub (50) – Cas (27) | Rub (50) – Lav (14) – Cas (12) |
| | Rub (50) – Ech (17) – Cas (16) | Rub (51) – Cas (11) |
| | Rub (52) – Eri (12) – Pru (11) | Rub (57) – Tri (22) |
| | Rub (58) – Aca (22) | Rub (61) – Ech (14) – Eri (13) |
| | Rub (62) – Ech (14) – Aca (13) | Rub (69) |
| <i>Trifolium</i> (≥45%) | Tri (58) – Cas (19) – Pru (14) | Tri (69) – Ech (10) |
| | Tri (82) – Euc (10) | |
| Multifloral | Cas (43) – Euc (27) | Cas (48) – Rub (19) – Eri (18) |
| | Cas (50) – Euc (18) – Tri (11) | Cas (54) – Rub (17) |
| | Cas (55) – Pru (20) – Rub (11) | Cas (58) – Pru (11) |
| | Cas (62) – Rub (10) | Cas (64) – Rub (14) |
| | Ech (39) – Rub (29) | Euc (44) – Ech (35) |
| | Euc (52) – Tri (33) | Euc (58) – Tri (15) – Cas (14) |
| | Pru (27) – Tri (23) – Cas (12) | Rub (39) – Ech (21) |
| | Tri (32) – Foe (24) – Lav (14) | |

*) Pólenes com predominância superiores a 10%.

Na Tabela 8 mostra-se o intervalo de cor na escala Pfund (mm) e a cor obtida, segundo as cores padrão do mel designadas pela USDA, de cada grupo de mel monofloral. Os intervalos de cor na escala Pfund (mm) para os méis classificados segundo os tipos apresentados na Tabela 7 são bastante alargados, mostrando que dentro do mesmo grupo existem várias cores padrão do mel. Por exemplo, os méis de *Lavandula* presentes neste trabalho (22 amostras) e os de *Echium* (17 amostras) têm cores extra-branco, branco, âmbar extra claro, âmbar claro e âmbar. De realçar que, os méis monoflorais escuros são principalmente dos tipos de pólen *Erica sp.* e *Castanea sp.*.

Nas amostras predominam os méis monoflorais (75 amostras). Destes, os mais representados são os méis monoflorais de *Lavandula sp.*, *Echium sp.*, *Rubus sp.* e *Erica sp.* (por ordem decrescente). Na amostragem estão também presentes méis monoflorais de *Prunus sp.*, *Castanea sp.*, *Trifolium sp.* e *Eucalyptus sp.*, mas com menor representação (4 ou menos amostras de méis). É evidente a variabilidade das amostras deste trabalho considerando os três tipos de pólen de cada amostra (só se apresentaram os que apresentavam quantidades superiores a 10%). A influência da variabilidade do perfil polínico de cada amostra pode ser verificada pela variação da cor dentro de cada grupo de amostras de mel monofloral.

Tabela 8 – Valores globais da classificação e da cor das amostras monofloral/multifloral.

| Mel | Tipo de pólen | Nº amostras | Cor (Pfund, mm) | Cor padrão * |
|-------------|-----------------------|-------------|-----------------|-----------------|
| Monofloral | <i>Echium sp.</i> | 17 | 15 - 138 | EB,B,AEC,AC,A,E |
| | <i>Erica sp.</i> | 8 | 126 - 201 | E |
| | <i>Eucalyptus sp.</i> | 2 | 23; 70 | B,AC |
| | <i>Castanea sp.</i> | 4 | 118 - 175 | E |
| | <i>Lavandula sp.</i> | 22 | 16 - 115 | EB,B,AEC,AC,A,E |
| | <i>Prunus sp.</i> | 4 | 61 - 156 | AC,E |
| | <i>Rubus sp.</i> | 14 | 40 - 196 | AEC,AC,A,E |
| | <i>Trifolium sp.</i> | 3 | 19; 66; 135 | B,AC,E |
| Multifloral | | 15 | 44 - 204 | |

*) EB - extra branco; B – branco; AEC - âmbar extra claro; AC - âmbar claro; A – âmbar; E – escuro

3.3. Análise por HPLC

Com o objetivo de verificar a qualidade das amostras de mel analisou-se o perfil de açúcares maioritários por HPLC. Pretendia-se determinar a concentração de três açúcares (sacarose, glucose e frutose) mas, nos cromatogramas das amostras, verificou-se que no pico da sacarose se sobrepunham outros dissacarídeos. Este facto foi verificado usando

soluções de padrões puros de turanose, trealose e maltose (dissacarídeos presentes no mel). Na Figura 6 mostra-se a sobreposição dos cromatogramas destes três padrões e de uma mostra de mel obtidos com uma coluna Supelcogel C- 610H (30 cm × 7,8 mm ID de dimensão) e um detetor IR. Verifica-se que o pico da sacarose da amostra de mel (tempo de retenção 12,5 min) está sobreposto com os picos dos 3 padrões. A glucose elui ao tempo de retenção 14,5 min e o de frutose, aos 15,8 min. Devido a esta sobreposição só foi possível quantificar a glucose e frutose nas amostras.

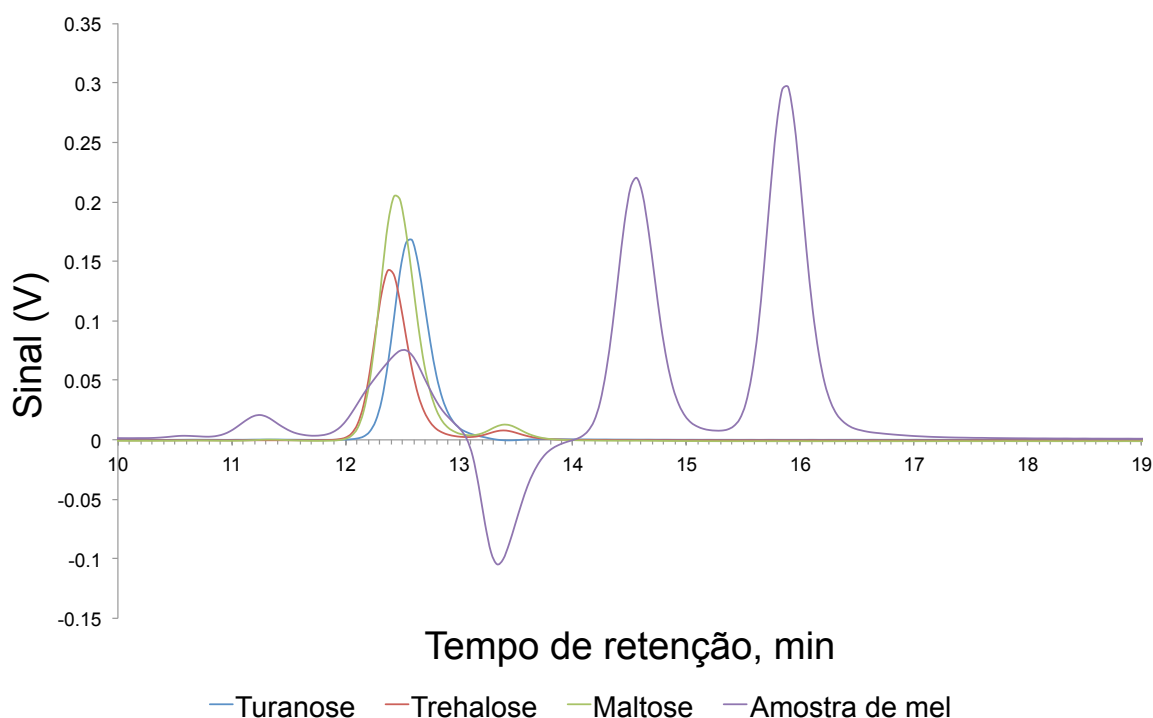


Figura 6 – Cromatogramas da turanose, trehalose, maltose e de uma mostra de mel.

A seguir, indicam-se os resultados obtidos da calibração do HPLC para a análise destes compostos usando soluções padrão de calibração, e da verificação da qualidade do método, feita com duas soluções de controlo de qualidade e três amostras de mel. A análise destes compostos foi efetuada através de curvas de calibração obtidas usando cinco soluções padrão preparadas por mistura da glucose e frutose com diferentes concentrações. Na Tabela 9 indicam-se os intervalos de concentrações usadas para cada composto nas soluções padrão de calibração e os parâmetros de cada calibração obtidos por regressão linear (declive, ordenada na origem e coeficiente de correlação), bem como os limites de deteção e quantificação calculados a partir destes. As curvas de calibração para a glucose e

frutose foram obtidas pela relação entre as áreas dos picos de cada composto em função das respetivas concentrações.

Tabela 9 – Intervalo de concentrações e parâmetros da calibração para a análise da glucose e frutose por HPLC.

| Composto | Cmin, g/L | Cmax, g/L | Declive ($\pm s$), L/g | OO ($\pm s$) | R | LD, g/L | LQ, g/L |
|----------|-----------|-----------|---------------------------|-------------------------|---------|---------|---------|
| Glucose | 0,50 | 4,97 | $(256 \pm 1) \times 10^4$ | $(4 \pm 4) \times 10^4$ | 0,99996 | 0,051 | 0,15 |
| Frutose | 0,59 | 4,89 | $(255 \pm 1) \times 10^4$ | $(2 \pm 4) \times 10^4$ | 0,99996 | 0,051 | 0,16 |

Cmin – concentração mínima; Cmax – concentração máxima; OO – ordenada na origem; R – coeficiente de correlação; LD – limite de deteção; LQ – limite de quantificação.

Obtiveram-se resultados semelhantes para a glucose e frutose: sensibilidades elevadas ($\approx 2,6 \times 10^6$ contagem \times L/g); coeficientes de correlação bons ($R \geq 0,9999$); limites de deteção baixos (0,05 g/L); e limites de quantificação baixos (0,15 g/L).

Para verificar o desempenho do método analítico estudou-se a precisão, ao nível da repetibilidade e precisão intermédia, e a exatidão, usando duas soluções de controlo de qualidade (uma de níveis de concentração baixos e outra, de concentrações médias). O estudo da precisão foi também efetuado usando três amostras de mel. Convém referir que as médias dos tempos de retenção das amostras e dos padrões são muito semelhantes e os valores de desvio padrão foram sempre inferiores a 0,03 min. Estas pequenas variações ($sr\% < 0,25\%$) mostram que a análise por HPLC usando o detetor IR, quer nos padrões quer nas amostras, é bastante reprodutível.

Na Tabela 10 apresentam-se os valores das concentrações médias e os desvios padrão associados e os desvios padrão relativos percentuais referentes ao estudo da repetibilidade e da precisão intermédia para as duas soluções de controlo de qualidade, bem como dos erros relativos percentuais considerando os valores esperados para cada composto. Globalmente, os resultados obtidos são aceitáveis ($sr\% \leq 5\%$). Mais, as concentrações medidas para cada composto nestas soluções de controlo de qualidade, no estudo da repetibilidade e da precisão intermédia, coincidiram. Os valores mais elevados foram obtidos com a solução de controlo de qualidade de concentrações médias, resultado que não era esperado. A glucose e frutose apresentaram valores aceitáveis de erro relativo percentual ($Er\% \leq 2\%$) mostrando que o método é exato.

Na Tabela 11 apresentam-se os resultados obtidos no estudo da repetibilidade e da precisão intermédia para três amostras, usando os valores de concentração obtidos com o detetor IR. As amostras apresentam níveis de concentração semelhantes de glucose e

frutose e globalmente, os resultados são aceitáveis ($sr\% \leq 5\%$) para as concentrações dos compostos analisados, ao nível da repetibilidade e precisão intermédia.

Tabela 10 – Resultados do estudo da repetibilidade, da precisão intermédia e da exatidão com soluções de controlo de qualidade.

| Solução | Composto | C _e , g/L | C _m ±s, g/L | Sr% | Er% |
|----------------------------|----------|----------------------|------------------------|-----|------|
| Repetibilidade | | | | | |
| SCQ1 | Glucose | 2,11 | 2,15±0,01 | 0,7 | 1,8 |
| | Frutose | 3,21 | 3,27±0,01 | 0,4 | 1,9 |
| SCQ2 | Glucose | 0,745 | 0,75±0,01 | 1,9 | 0,2 |
| | Frutose | 0,731 | 0,74±0,01 | 1,6 | 0,8 |
| Precisão intermédia | | | | | |
| SCQ1 | Glucose | 2,11 | 2,1±0,1 | 4,9 | 0,7 |
| | Frutose | 3,21 | 3,2±0,1 | 3,4 | 0,7 |
| SCQ2 | Glucose | 0,745 | 0,74±0,01 | 1,7 | 0,08 |
| | Frutose | 0,731 | 0,74±0,01 | 1,7 | 0,8 |

SCQ – solução de controlo de qualidade; C_e – concentração esperada; C_m – concentração medida; s – desvio padrão; sr% - desvio padrão relativo percentual; Er% - erro relativo percentual

Tabela 11 – Resultados do estudo da repetibilidade e da precisão intermédia de 3 amostras.

| Solução | Composto | C _m ±s, g/100g mel | Sr% |
|----------------------------|----------|-------------------------------|-----|
| Repetibilidade | | | |
| Amostra 1 | Glucose | 25,6±0,2 | 0,7 |
| | Frutose | 39,0±0,2 | 0,4 |
| Amostra 2 | Glucose | 25,22±0,09 | 0,3 |
| | Frutose | 42,81±0,08 | 0,2 |
| Amostra 3 | Glucose | 27,7±0,1 | 0,4 |
| | Frutose | 39,20±0,09 | 0,2 |
| Precisão intermédia | | | |
| Amostra 1 | Glucose | 25±1 | 5,0 |
| | Frutose | 39±1 | 3,4 |
| Amostra 2 | Glucose | 25,6±0,3 | 1,0 |
| | Frutose | 43,3±0,5 | 1,0 |
| Amostra 3 | Glucose | 28,5±0,6 | 2,1 |
| | Frutose | 39,7±0,3 | 0,7 |

C_m – concentração média; s – desvio padrão; sr% - desvio padrão relativo percentual

Em suma, os resultados do estudo de desempenho do método usando a coluna de exclusão molecular mostraram que este era adequado para a análise das amostras de mel, com vista a averiguar a variabilidade dos níveis de concentração da glucose e frutose nas amostras (açúcares maioritários). Na Tabela 12 mostram-se os resultados globais obtidos na análise de glucose e frutose nas amostras de mel por HPLC, bem como do cálculo da concentração dos monossacarídeos (por adição das concentrações daqueles dois açúcares maioritários) e da relação entre as concentrações de frutose e glucose.

Em geral, a frutose está presente com uma concentração mais elevada mas, em algumas amostras, a glucose pode ser o açúcar maioritário. Os níveis de frutose nas amostras variaram entre 33 e 57 g/100 g de mel e os de glucose entre 20 e 37 g/100 g de mel. A quantidade de monossacarídeos nas amostras, calculada pela adição das concentrações da glucose e frutose, varia entre 54 e 90 g/100 g de mel. Para a razão entre as concentrações de frutose e glucose, os resultados foram sempre superiores a 1, variando entre 1,2 e 1,7.

Tabela 12 – Concentrações de glucose e frutose nas amostras de mel analisadas por HPLC

| Parâmetro | Cmin, g/100g mel | Cmax, g/100g mel | C _m ±s, g/100g mel |
|---------------------------|------------------|------------------|-------------------------------|
| Glucose | 20,6 | 37,1 | 28±4 |
| Frutose | 33,3 | 57,0 | 42±5 |
| Monossacrídeos | 53,9 | 90,0 | 70±8 |
| | Rmin | Rmax | R_m±s |
| Razão [frutose]/[glucose] | 1,16 | 1,74 | 1,5±0,1 |

Cmin – concentração mínima; Cmax – concentração máxima; C_m – concentração média; s – desvio padrão; Monossacrídeos = Glucose + frutose; Rmin – razão mínima; Rmax – razão máxima; R_m – razão média

Os níveis apresentados dos açúcares estão de acordo, em parte, com os resultados obtidos em outros trabalhos [3,51] que referem que, em geral, no mel a percentagem de frutose varia entre 30 e 45% e a de glucose, entre 22 e 41%. Comparando os valores de concentração medidos verifica-se que, em algumas amostras obtiveram-se valores superiores a 45% para a concentração da frutose e inferiores a 22%, para a glucose. Estes resultados podem estar relacionados com os valores de % de humidade presentes nas amostras (parâmetro não analisado).

Estes resultados evidenciam uma grande variabilidade na composição das amostras em açúcares (os parâmetros apresentados na Tabela 12 mostram desvios padrão relativos percentuais na ordem dos 10-12%), que se deverá ter presente na discussão dos resultados

da discriminação dos méis monovarietais através dos perfis dos sinais obtidos da análise com a LE.

3.4. Análise com a língua eletrónica

Na secção 3.2, através da análise polínica, verificou-se que das 89 amostras usadas, 75 amostras correspondem a méis monoflorais e somente 14 a méis multiflorais. Nas amostras estão presentes 8 tipos de méis monoflorais: *Lavandula sp.*, *Echium sp.*, *Rubus sp.* e *Erica sp.* (mais representados); *Prunus sp.*, *Castanea sp.*, *Trifolium sp.* e *Eucalyptus sp.* (menos representados, ≤ 4 méis).

Todas as amostras foram analisadas com a LE e os sinais de potencial (V) de cada sensor estão representados na Figura 7. Os perfis dos sinais da LE obtidos nestas análises mostram, em geral, resultados semelhantes, o que indica que os sensores responderam de forma semelhante para todas as amostras.

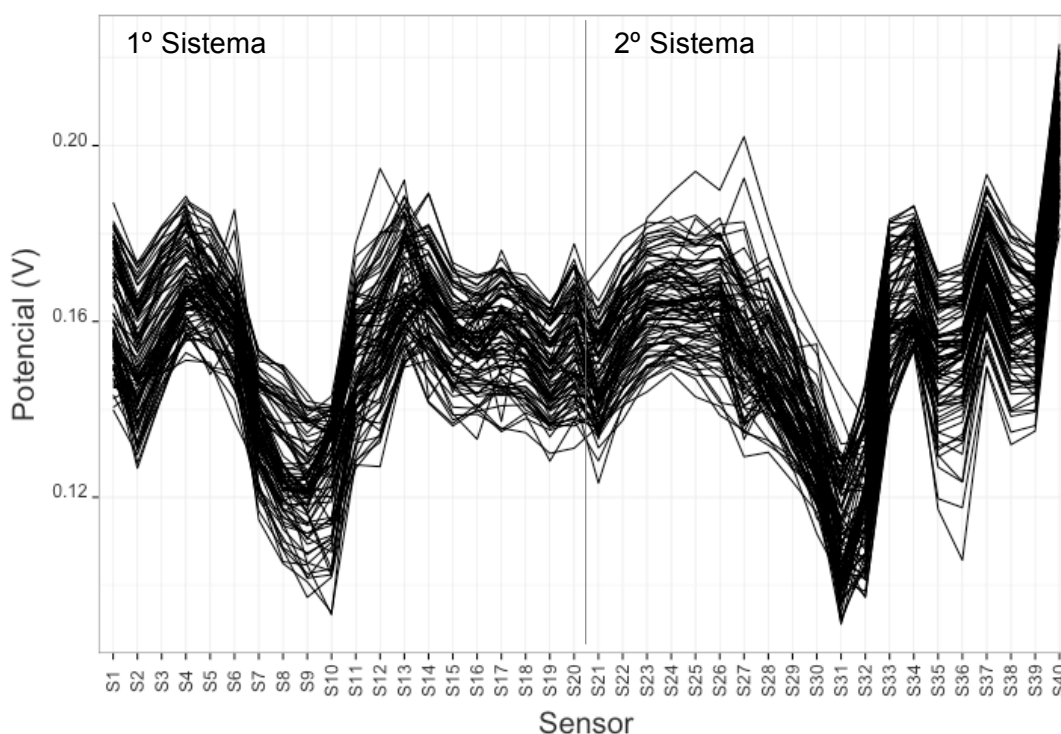


Figura 7 – Perfil dos sinais obtidos da análise das amostras com a língua eletrónica (S21 a S40, duplicados das membranas S1 a S20, respetivamente).

A Figura 8 mostra o gráfico *boxplot* (gráfico de extremos quartis) onde se visualiza a variação dos sinais em cada sensor. Este gráfico mostra a presença de amostras com valores extremos em 9 dos sensores (sensores S5, S6, S10, S12, S25, S27, S30, S35 e

S36), mas as amostras identificadas não são comuns, ou seja, são valores de sinal máximos ou mínimos para diferentes amostras, indicando tratar-se da resposta do sensor à matriz da amostra em questão. A amostra 15 foi extremo em 2 sensores, a amostra 23 em 3 sensores, a amostra 28 em 4 amostras, a amostra 45 em 2 amostras, a amostra 61 em 1 sensor e a amostra 78 em 4 sensores. Por isso, estas amostras não foram removidas nos tratamentos de dados subsequentes.

Como referido, neste estudo pretende-se investigar se a LE permite classificar méis de acordo com a classificação polínica, necessitando de estabelecer-se uma metodologia de tratamento de dados para atingir este objetivo. A técnica supervisionada de classificação aplicada é a ADL que tem menor desempenho de previsão quando as variáveis independentes são multicolineares (em geral, com $R > 0,8$), como os sinais dos sensores da LE [52]. Quando existe elevada multicolinearidade entre variáveis, o modelo deve ser avaliado atendendo à sua robustez ao nível das previsões, pelo que a comparação dos modelos será feita, tendo em conta as classificações corretas das amostras considerando os dados originais e, principalmente, as obtidas usando a validação cruzada LOO. A técnica da validação cruzada é importante pois permite evitar uma modulação de dados demasiado optimista.

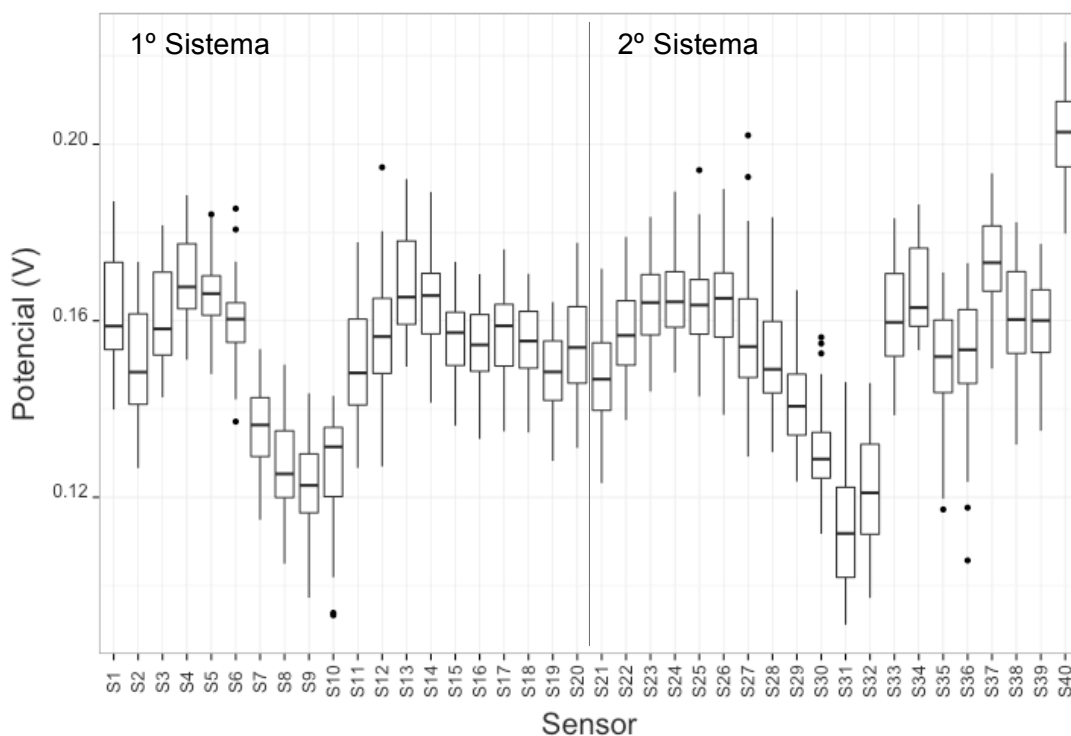


Figura 8 – Extremos quartis do potencial (mV) em função dos sensores (S21 a S40, duplicados das membranas S1 a S20, respetivamente).

Neste âmbito, a questão mais relevante que se coloca em sistemas de multi-parâmetros é a seleção de variáveis. O uso de todas as variáveis pode produzir um bom modelo de ajuste para os dados originais, porque os resíduos são menores e, por isso, o coeficiente de correlação maior. Mas, o interesse não é maximizar o ajuste para os dados de treino (dados originais) mas, o desempenho de previsão para os dados do teste que é avaliado pela validação cruzada LOO. Assim, uma redução das variáveis independentes pode evitar os efeitos de sobre-ajustamento entre a variável dependente e as variáveis independentes e levar a um melhor desempenho de previsão [53]. Para este efeito, estudou-se a aplicação de 3 algoritmos meta-heurísticos de seleção de variáveis para comparação e seleção do que melhor se adequa aos dados da LE: algoritmo arrefecimento simulado (AS), algoritmo genético (GEN) e algoritmo melhoramento local restringido (MLR). Na aplicação de cada um destes métodos usou-se o critério ccr12 para avaliar a qualidade do nível de ajuste entre variável dependente (grupos) e as variáveis independentes (subconjuntos de sensores) selecionadas. Para cada algoritmo, o estudo envolveu a procura do melhor modelo para cada um dos subconjuntos definidos (neste trabalho, foi de 2 a 20 sensores). Os 19 modelos obtidos de cada algoritmo usado, posteriormente, são testados na sua capacidade de classificação dos méis monoflorais (quer com dados originais quer com a metodologia validação cruzada LOO), selecionando-se o que mostrava melhores resultados de classificações corretas para a validação cruzada LOO. Na Figura 9 mostra-se um esquema que resume este procedimento de tratamento de dados.

O primeiro estudo das capacidades analíticas da LE para classificar méis monoflorais incidiu, numa primeira aproximação, na discriminação simultânea de todos os tipos de méis monoflorais presentes, sem considerar o grupo de méis multiflorais, através da ADL. Esta primeira análise dos dados analíticos implicou a discriminação de 8 grupos de méis: *Echium* (Ech), *Erica* (Eri), *Eucalyptus* (Euc), *Castanea* (Cas), *Lavandula* (Lav), *Prunus* (Pru), *Rubus* (Rub) e *Trifolium* (Tri). A utilização dos algoritmos de seleção de variáveis não permitiu obter qualquer modelo discriminante sub-conjunto dos 40 sensores iniciais com bons resultados de previsão. O modelo a seguir apresentado (Figura 10) corresponde à utilização de todos os 40 sensores, que permitiu obter percentagens de classificações corretas de 95,9% e 28,4%, respectivamente, para os dados originais e para a validação cruzada LOO. No modelo ADL obtiveram-se 7 funções discriminantes que, por ordem crescente do número da função, explicam 43, 18, 14, 9, 7, 5 e 4% da variabilidade dos dados. Verifica-se também que a inclusão de grupos de tipo de mel representadas com apenas 2 amostras (mel *Eucalyptus* e *Trifolium*) não contribui para a obtenção de modelos robustos, pois haverá situações em que, ao aplicar a validação cruzada LOO, o modelo estudado têm pouca informação sobre esse tipo de mel monofloral contribuindo para uma diminuição da percen-

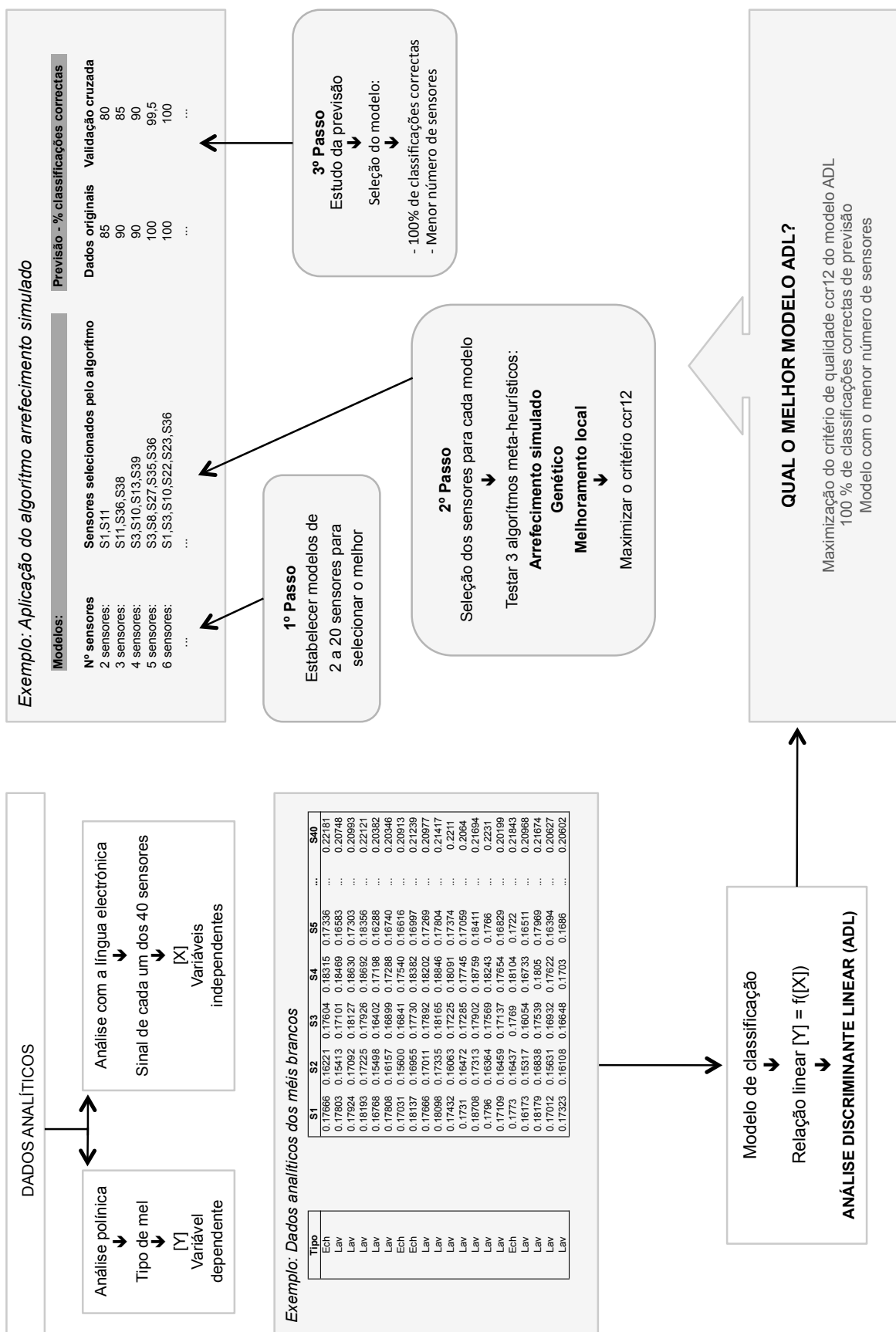


Figura 9 – Esquema da metodologia de tratamento de dados usada para estabelecer o melhor modelo ADL.

tagem de classificações corretas, mesmo ao nível dos dados originais. No caso do grupo dos méis de *Eucaliptus*, presente com duas amostras, a figura mostra a presença de uma reta que corresponde à elipse associada ao intervalo de confiança a 95%.

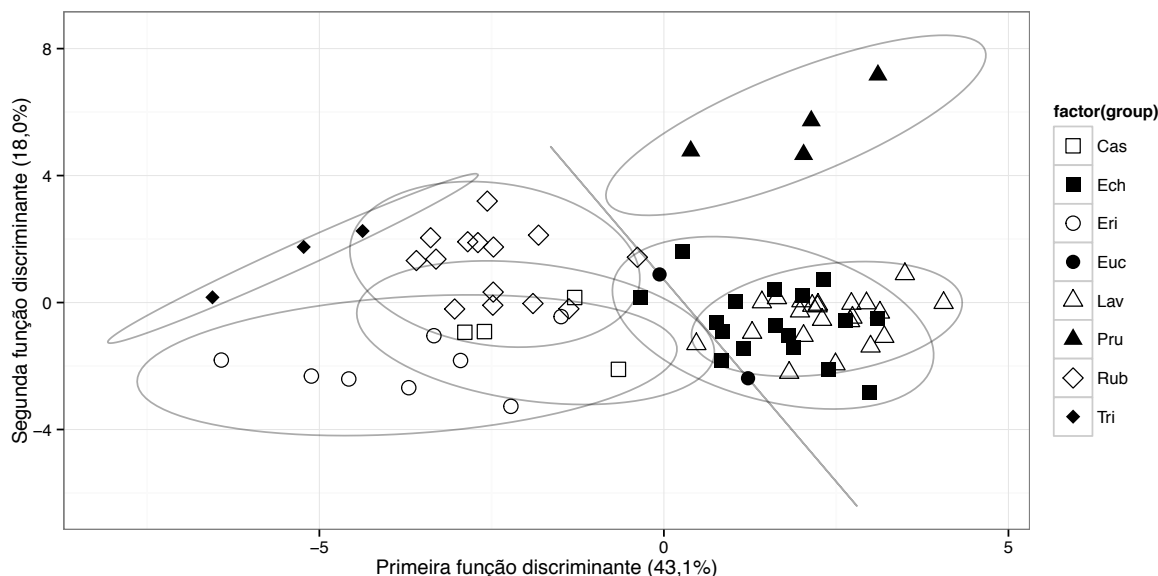


Figura 10 – Análise discriminante linear entre 8 diferentes tipos de mel. As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

Na Figura 11 mostra-se, usando os 40 sensores, a discriminação dos 6 grupos de tipos de méis com representação de 3 ou mais amostras. O modelo de discriminação obtido para os 6 tipos de mel (Ech, Eri, Cas, Lav, Pru e Rub) consistiu em 5 funções discriminantes cuja percentagem de variância explicada por cada função é de 52, 19, 16, 7 e 5%. Reduziu-se o número de funções e as primeiras explicam uma maior variância dos dados, contribuindo para o aumento das percentagens de classificações corretas para os dados originais e para a validação cruzada: 96,0 e 44,6 %, respetivamente. Esta figura mostra que há uma grande intra-variabilidade nas amostras pertencentes ao mesmo tipo de mel monofloral, o que dificulta a discriminação entre os diferentes tipos de mel. Verifica-se também que entre os méis de *Lavandula* e *Echium* há sobreposição de amostras (méis mais claros), tal como entre as amostras de méis de *Erica*, *Rubus* e *Castanea* (méis mais escuros). Os méis de *Prunus* têm perfis de sinais bastante distintos em relação aos outros tipos de mel. Esta situação mostrou que a cor poderia ser um parâmetro de seleção de amostras para simplificar o tratamento estatístico dos dados analíticos de forma a obter uma metodologia multivariada para atingir classificações corretas.

A solução encontrada para simplificar este problema foi fazer uma seleção primária das amostras em função de 3 cores definidas da seguinte maneira: branco (valores ≤ 34 mm Pfund, inclui as cores de branco água, extra branco e branco), âmbar (valores > 34 e ≤ 50 mm Pfund, inclui as cores âmbar extra claro, âmbar claro e âmbar) e escuro (valores > 114 mm Pfund, correspondendo unicamente à cor de âmbar escuro). Assim, por exemplo, as amostras de *Lavandula* poderão ser identificadas em dois grupos de cor de mel (branco e âmbar) e, devido a haver menor número de tipos de mel dentro de cada um dos grupos de cor, a classificação será mais robusta.

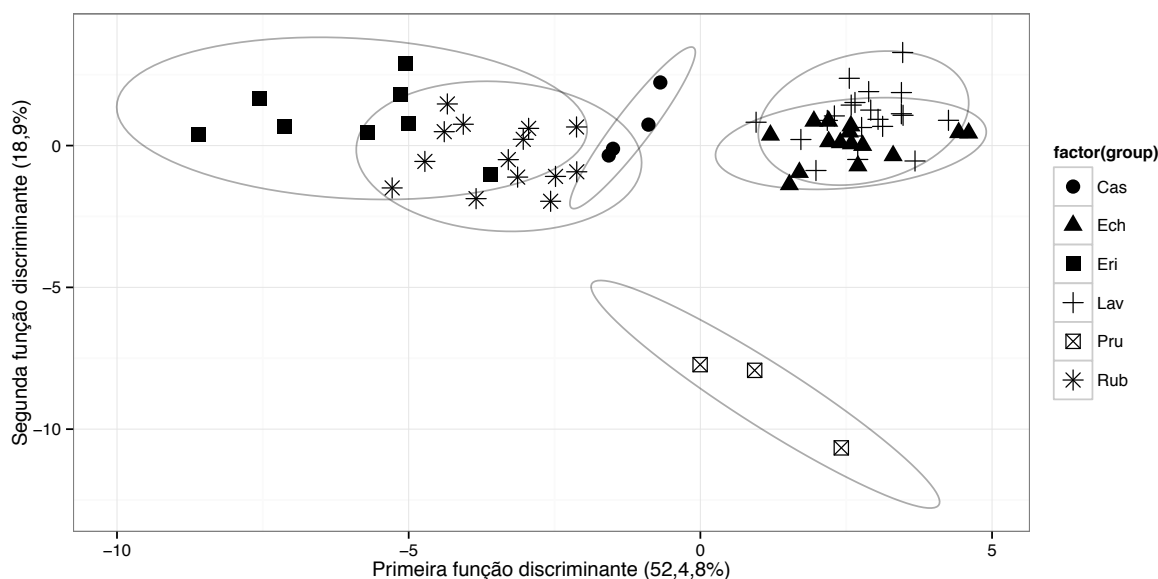


Figura 11 – Análise discriminante linear entre 6 diferentes tipos de mel. As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

Para verificar se há informação nos perfis de sinais da análise das amostras com a LE segundo estes 3 grupos de cor, efetuou-se ADL usando os sinais dos 40 sensores referentes às amostras classificadas como méis monoflorais. Na Figura 12 mostra-se a representação gráfica das 2 primeiras funções discriminantes obtidas (explicam 100% da variância presente nos dados). Esta figura mostra a distribuição das amostras em relação às duas funções discriminantes obtidas para o modelo com os dados originais que permite 100% de classificações corretas em relação aos 3 grupos de cor definidos. Verifica-se que os perfis de sinais da LE permitem esta separação, o que é também indicativo da validade desta estratégia a nível laboratorial e no tratamento de dados.

Os estudos seguintes tiveram como objetivo verificar se a LE permite classificar amostras de mel monofloral corretamente. Devido à complexidade das amostras que se

reflete nos perfis de sinais obtidos da análise com a LE, decidiu-se investigar se a análise laboratorial das amostras poderia ser efetuada tendo em conta a sua separação em três grupos de cor: branco, âmbar e escuro. Este procedimento poderia permitir uma possível aplicação prática desta ferramenta na análise de amostras de mel desconhecidas, cuja eficácia dependerá da existência ou não de padrões de mel adequados para cada grupo de mel definido. Para demonstrar a viabilidade deste procedimento, apresentam-se a seguir três estudos de análise discriminante para a classificação de méis monoflorais de cor branca, âmbar e escura.

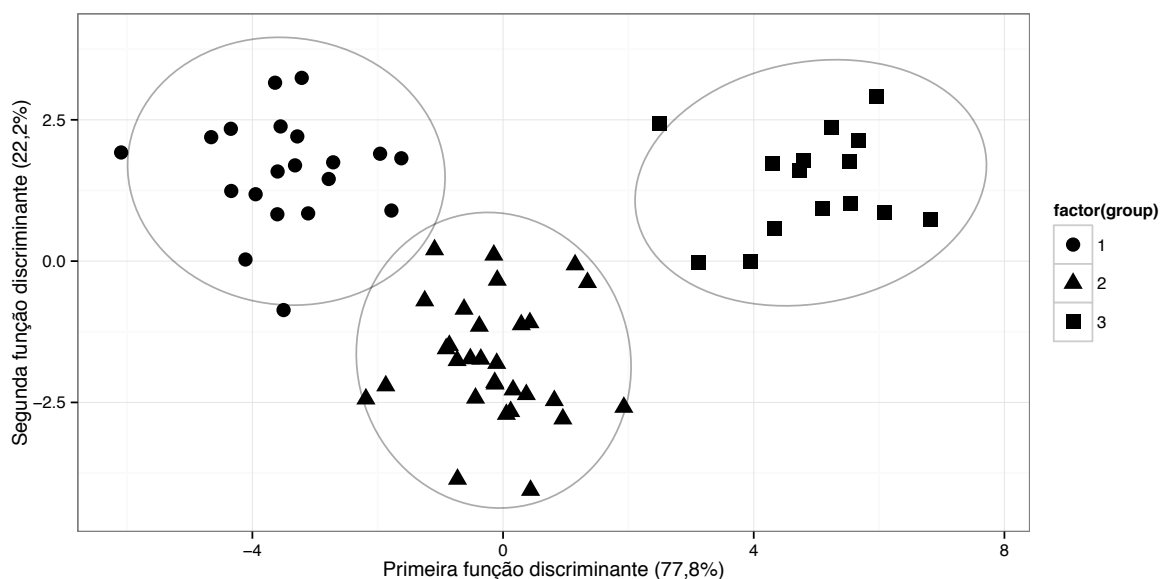


Figura 12 – Análise discriminante linear entre os 3 grupos de cor associados à seleção primária das amostras: 1 – branco; 2 – âmbar; 3 – escuro. As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

3.4.1. Classificação de amostras de cor branca

No grupo das amostras de mel branco (amostras de cor extra branco e branco) incluem-se dois tipos de mel monofloral: *Echium* (4 amostras), *Lavandula* (16 amostras). As amostras únicas de *Eucalyptus* e *Trifolium* foram excluídas, por não cumprirem o critério de os grupos conterem um mínimo de três amostras.

Na Tabela 13 apresentam-se os resultados de aplicação dos 3 algoritmos para a seleção de variáveis independentes (sensores) de modo a obter modelos discriminantes com um número variável de sensores (2 a 20 sensores) e permitir verificar qual o modelo mais simples e eficiente na previsão do tipo de mel monofloral de cor branca. Para cada modelo obtido apresentam-se os valores do critério de qualidade ccr12 associado ao modelo

discriminante dos dois grupos definidos em função das variáveis independentes selecionadas, bem como das percentagens de classificação corretas quer para os dados originais quer para a validação cruzada LOO. Na tabela mostram-se todos os modelos obtidos para o algoritmo AS e os obtidos para os algoritmos GEN e MLR que foram diferentes destes.

Tabela 13 – Seleção de variáveis independentes através de 3 algoritmos para obter modelos de discriminação entre 2 grupos de mel monofloral de cor branco.

| Nº sensores | Sensores selecionados | ccr12 | %CC DO | %CC LOO |
|----------------------|--|------------------|------------|------------|
| Algoritmo AS | | | | |
| 2 | 1,11 | 0.4246743 | 85 | 80 |
| 3 | 11,36,38 | 0.5980923 | 90 | 85 |
| 4 | 3,10,13,39 | 0.6970949 | 90 | 90 |
| 5 | 3,8,27,35,36 | 0.7985231 | 100 | 99,5 |
| 6 | 1,3,10,22,23,36 | 0.9310705 | 100 | 100 |
| 7 | 1,3,4,10,22,23,36 | 0.9742734 | 100 | 100 |
| 8 | 1,3,4,10,14,25,30,36 | 0.9867681 | 100 | 100 |
| 9 | 1,3,4,8,10,14,25,30,36 | 0.9938599 | 100 | 100 |
| 10 | 1,3,4,10,15,22,23,25,30,36 | 0.9977131 | 100 | 100 |
| 11 | 1,3,4,10,15,20,22,23,25,30,36 | 0.9990685 | 100 | 100 |
| 12 | 1,3,8,10,11,19,22,23,27,35,36,37 | 0.9996069 | 100 | 100 |
| 13 | 1,3,7,8,10,11,19,22,23,27,35,36,37 | 0.9998974 | 100 | 100 |
| 14 | 1,3,4,8,10,11,18,19,22,23,27,35,36,37 | 0.9999905 | 100 | 100 |
| 15 | 1,3,4,8,9,10,14,16,19,24,25,29,30,31,36 | 0.9999997 | 95 | 60 * |
| 16 | 1,3,4,8,10,11,14,18,19,22,23,30,33,36,37,39 | 1.0000000 | 100 | 60 * |
| 17 | 1,2,3,7,8,12,13,16,22,24,26,28,31,33,34,36,37 | 1.0000000 | 100 | 50 * |
| 18 | 1,2,4,7,8,9,10,12,17,20,22,24,25,26,31,33,35,37 | 1.0000000 | 100 | 45 * |
| 19 | 1,4,9,11,12,15,16,17,19,21,23,24,29,30,32,34,37,38,40 | 1.0000181 | 90 | 55 * |
| 20 | ** | | | |
| Algoritmo GEN | | | | |
| | | | | |
| 15 | 1,3,10,11,14,19,22,23,27,30,33,35,36,37,39 | 0.9999981 | 100 | 95 |
| 16 | | | 100 | 60 * |
| 17 | 1,3,4,8,10,11,14,18,19,22,23,27,33,34,35,36,37 | 1.0000000 | 100 | 50 * |
| 18 | 1,2,3,5,6,8,9,12,18,25,27,29,31,33,36,38,39,40 | 1.0000000 | 100 | 45 * |
| 19 | 3,5,11,12,13,14,17,18,23,24,26,27,29,30,31,33,34,35,38 | 1.0000153 | 100 | 40 * |
| 20 | ** | | | |
| Algoritmo MLR | | | | |
| | | | | |
| 12 | 1,3,7,12,15,18,22,23,24,31,36,39 | 0.9994526 | 100 | 100 |
| 13 | | | | |
| 14 | 1,3,8,10,11,14,19,22,23,30,33,36,37,39 | 0.9999722 | 100 | 100 |
| 15 | 1,2,3,5,6,10,12,13,18,22,23,28,35,36,37 | 0.9999975 | 100 | 100 |
| 16 | 3,5,8,10,11,17,22,23,24,25,26,27,31,36,38,39 | 1.0000000 | 100 | 80 * |
| 17 | 1,2,3,10,13,15,16,21,22,24,26,29,30,31,32,35,39 | 1.0000000 | 95 | 45 * |
| 18 | 1,3,4,7,8,10,11,13,17,21,22,23,27,29,30,37,38,40 | 1.0000000 | 100 | 35 * |
| 19 | 1,2,7,10,12,17,19,20,23,24,27,29,30,31,32,34,35,37,38 | 1.0000107 | 95 | 75 * |
| 20 | ** | | | |

%CC DO – Percentagem de classificações corretas com os dados originais;

%CC LOO – Percentagem de classificações corretas com validação cruzada LOO;

*) Problemas de colinearidade verificadas na validação cruzada LOO;

**) O algoritmo não foi capaz de encontrar qualquer subconjunto não-colinear confiável de dimensionalidade 20.

Globalmente, os 3 algoritmos permitiram obter modelos com sub-conjuntos de 2 a 19 sensores usando os dados originais. Para o modelo com 20 sensores, os 3 algoritmos indicaram não terem conseguido um sub-conjunto de sensores fiável, o que indica a ocorrência de problemas de colinearidade. Problemas de colinearidade foram também encontrados a partir dos modelos com 15 ou 16 sensores, ao nível da validação cruzada LOO, para os algoritmos utilizados. Os resultados mostram também que para a discriminação entre os dois grupos de méis de *Echium* e *Lavandula* há um número ideal de sensores, neste caso entre 6 e 14 sensores, que permitem 100% de classificações corretas na validação cruzada LOO. Verifica-se que o critério de qualidade (ccr12) do modelo obtido aumenta com o número de sensores, mas um valor de ccr12 >0,99 não é indicativo de um modelo com 100% de classificações corretas principalmente na validação cruzada LOO. Quando se acrescentam sensores ao modelo com 14 sensores, a inserção faz diminuir a percentagem das classificações corretas na validação cruzada LOO e, em alguns casos, também no modelo com os dados originais. Esta situação mostra que nem sempre o aumento de variáveis num modelo de discriminação implica um aumento do seu desempenho de previsão (e em certos casos, de estimação).

Os três algoritmos testados chegaram ao mesmo modelo que apresentou menor número de sensores selecionados (modelo com 7 sensores), maior valor de critério de qualidade ccr12 e 100% de classificações corretas na validação cruzada, usando o método LOO, entre os dois grupos de méis brancos de *Echium* e *Lavandula*.

A Figura 13 mostra a representação gráfica da distribuição de frequência dos dados referentes à primeira e única função discriminante obtida (explica 100% da variância dos dados), sendo possível verificar que os dois grupos de dados originais estão bem discriminados, tal como os resultados de previsão indicavam.

3.4.2. Classificação de amostras de cor âmbar

Nas amostras de cor âmbar (amostras de cor âmbar extra claro, âmbar claro e âmbar) obtiveram-se 4 grupos de méis monoflorais: *Echium* (11 amostras), *Lavandula* (5 amostras), *Prunus* (3 amostras) e *Rubus* (11 amostras). Tal como no exemplo anterior, pretendeu-se verificar se o perfil de sinais da LE permite distinguir estes 4 grupos de amostras comparando modelos obtidos da aplicação dos 3 algoritmos de seleção de variáveis presentemente em análise. Neste estudo retiraram-se as amostras de mel monofloral de *Trifolium* e de *Eucalyptus* por só haver uma amostra de cada.

Na Tabela 14 apresentam-se os resultados de aplicação dos 3 algoritmos para a seleção de sub-conjuntos de sensores com vista a obter modelos de discriminação linear (cada modelo contém um número diferente de sensores, de 2 a 20 sensores) dos 4 tipos de

mel monofloral de cor âmbar acima referidos. A estrutura desta Tabela é igual à da anterior. No caso do modelo obtido com o algoritmo MLR ser igual ao obtido com o algoritmo GEN, assinala-se à frente da indicação do número de sensores no modelo.

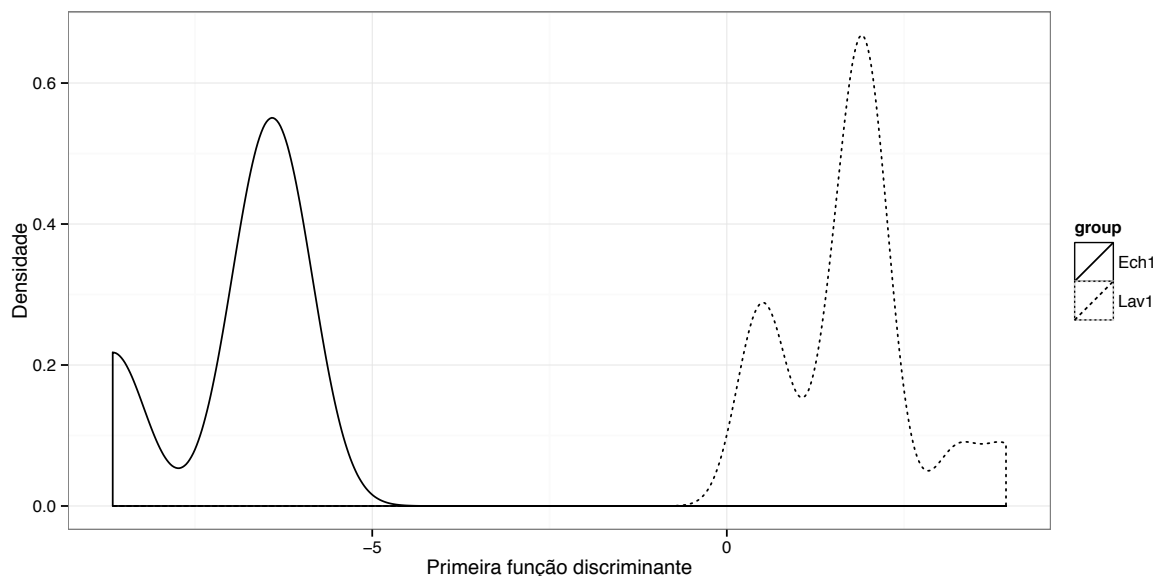


Figura 13 – Distribuição de frequências dos dados referentes à primeira função discriminante da classificação de dois grupos de méis brancos monoflorais: *Echium* (Ech) e *Lavandula* (Lav).

Os 3 algoritmos permitiram obter todos os modelos definidos e sem indicação de qualquer problema ao nível da colinearidade entre os sensores. O modelo com menor número de sensores selecionados (16 sensores) e 100% de classificações corretas na validação cruzada LOO foi obtido pelo algoritmo AS. Os modelos subsequentes mostraram também terem desempenhos de previsão de classificações 100% corretas. O algoritmo GEN forneceu só um modelo com este desempenho, com 20 sensores selecionados. No caso do algoritmo MLR, os modelos com sub-conjuntos de 18 a 20 sensores permitem 100% de classificações corretas na validação cruzada LOO.

O modelo selecionado apresenta três funções discriminantes que explicam 98,56%, 1,11% e 0,32% da variância total dos dados originais. Na Figura 14 mostra-se a representação gráfica das duas primeiras funções discriminantes associadas à separação dos 4 grupos de méis monoflorais e as elipses de cada grupo associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

Verificou-se, como era esperado, a separação entre os quatro grupos de méis monoflorais de cor âmbar e que a primeira função (explica 98% da variância dos dados) tem informação necessária para este efeito. Segundo esta dimensão, as amostras de *Rubus* são

Tabela 14 – Seleção de variáveis independentes através de 3 algoritmos para obter modelos de discriminação entre 4 grupos de mel monofloral de cor âmbar.

| Nº sensores | Sensores selecionados | ccr12 | %CC DO | %CC LOO |
|----------------------|--|------------------|------------|------------|
| Algoritmo AS | | | | |
| 2 | 14,21 | 0.6849298 | 70 | 67 |
| 3 | 7,11,37 | 0.7465714 | 77 | 57 |
| 4 | 7,17,24,37 | 0.7958008 | 97 | 87 |
| 5 | 7,11,16,27,37 | 0.8474409 | 93 | 73 |
| 6 | 7,11,15,22,24,37 | 0.8656918 | 97 | 83 |
| 7 | 14,17,21,25,34,36,39 | 0.8958645 | 97 | 87 |
| 8 | 5,7,9,11,18,21,23,40 | 0.9401209 | 97 | 73 |
| 9 | 2,14,17,23,28,36,37,38,39 | 0.9571124 | 100 | 97 |
| 10 | 2,14,17,23,26,28,36,37,38,39 | 0.9604490 | 100 | 97 |
| 11 | 2,3,6,7,9,11,18,21,23,38,40 | 0.9736348 | 100 | 90 |
| 12 | 2,3,7,9,11,14,20,22,23,32,39,40 | 0.9805893 | 97 | 90 |
| 13 | 2,3,4,8,10,11,12,19,20,23,27,30,33 | 0.9851214 | 93 | 87 |
| 14 | 2,3,7,11,13,19,20,21,22,25,30,32,33,40 | 0.9917983 | 100 | 63 |
| 15 | 8,13,14,16,17,18,21,23,24,29,30,31,35,36,37 | 0.9955507 | 100 | 97 |
| 16 | 6,7,8,10,12,14,17,21,24,28,32,34,35,36,38,39 | 0.9979479 | 100 | 100 |
| 17 | 6,7,8,10,12,14,17,21,24,28,32,34,35,36,38,39,40 | 0.9985504 | 100 | 100 |
| 18 | 2,3,5,7,9,11,12,15,19,20,22,24,26,27,30,33,38,40 | 0.9994228 | 100 | 100 |
| 19 | 1,3,5,6,7,13,15,16,17,18,19,20,21,29,31,32,34,36,38 | 0.9998251 | 100 | 100 |
| 20 | 6,7,8,10,12,13,14,17,19,21,24,28,30,31,32,34,35,36,38,39 | 0.9999109 | 100 | 100 |
| Algoritmo GEN | | | | |
| | | | | |
| 7 | 7,11,12,24,33,37,38 | 0.8853559 | 93 | 83 |
| 8 | | | | |
| 9 | 5,7,9,11,14,18,21,23,40 | 0.9490175 | 93 | 87 |
| 10 | 5,7,9,10,11,12,14,21,23,40 | 0.9582284 | 97 | 73 |
| 11 | | | | |
| 12 | 2,3,6,7,9,11,16,18,21,23,38,40 | 0.9792950 | 100 | 93 |
| 13 | 2,3,6,7,9,11,16,18,21,22,23,38,40 | 0.9834889 | 100 | 87 |
| 14 | 2,3,7,9,11,18,19,20,21,22,23,25,32,40 | 0.9864985 | 97 | 60 |
| 15 | 2,3,7,9,11,18,19,20,21,22,23,25,31,32,40 | 0.9893943 | 100 | 73 |
| 16 | 2,3,5,6,7,9,11,16,18,20,22,23,25,32,38,40 | 0.9921101 | 100 | 90 |
| 17 | 2,3,7,9,11,12,14,20,22,23,31,32,35,36,37,39,40 | 0.9947790 | 97 | 83 |
| 18 | 2,3,6,7,9,10,11,12,13,19,20,22,25,30,32,33,35,40 | 0.9986919 | 100 | 77 |
| 19 | 2,3,6,7,9,11,13,16,18,19,20,22,25,27,30,32,36,37,40 | 0.9995156 | 100 | 90 |
| 20 | 1,3,4,5,10,13,14,16,17,18,21,23,24,29,31,32,35,36,37,38 | 0.9998292 | 100 | 100 |
| Algoritmo MLR | | | | |
| | | | | |
| 12 = GEN | | | | |
| 13 | 2,3,7,9,11,14,18,20,22,23,32,39,40 | 0.9834889 | 96,7 | 83 |
| 14 | | 0.9864985 | 100 | 63 |
| 15 | | 0.9893943 | 100 | 93 |
| 16 | 4,10,13,14,16,17,19,21,22,23,24,29,31,35,37,38 | 0.9921101 | 100 | 97 |
| 17 | 3,6,7,8,10,12,14,17,21,24,28,32,34,35,36,38,39 | 0.9947790 | 100 | 97 |
| 18 | | 0.9986919 | 100 | 100 |
| 19 | 3,4,5,6,7,14,16,17,23,24,25,26,29,30,32,33,35,37,38 | 0.9995156 | 100 | 100 |
| 20 | 1,2,3,4,5,6,7,13,14,16,17,21,22,23,24,29,30,35,37,38 | 0.9998292 | 100 | 100 |

%CC DO – Percentagem de classificações corretas com os dados originais

%CC LOO – Percentagem de classificações corretas com validação cruzada LOO

*) Problemas de colinearidade verificadas na validação cruzada LOO

as que mostram perfis de sinais de sensores mais distintos (conjunto de amostras no quadrante positivo). As amostras de *Prunus* e *Echium* estão no quadrante negativo desta

função discriminante mostrando perfis de sinais semelhantes, enquanto os méis de *Lavandula* se encontram na zona central, indicativo de serem amostras com características intermédias em relação aos outros grupos.

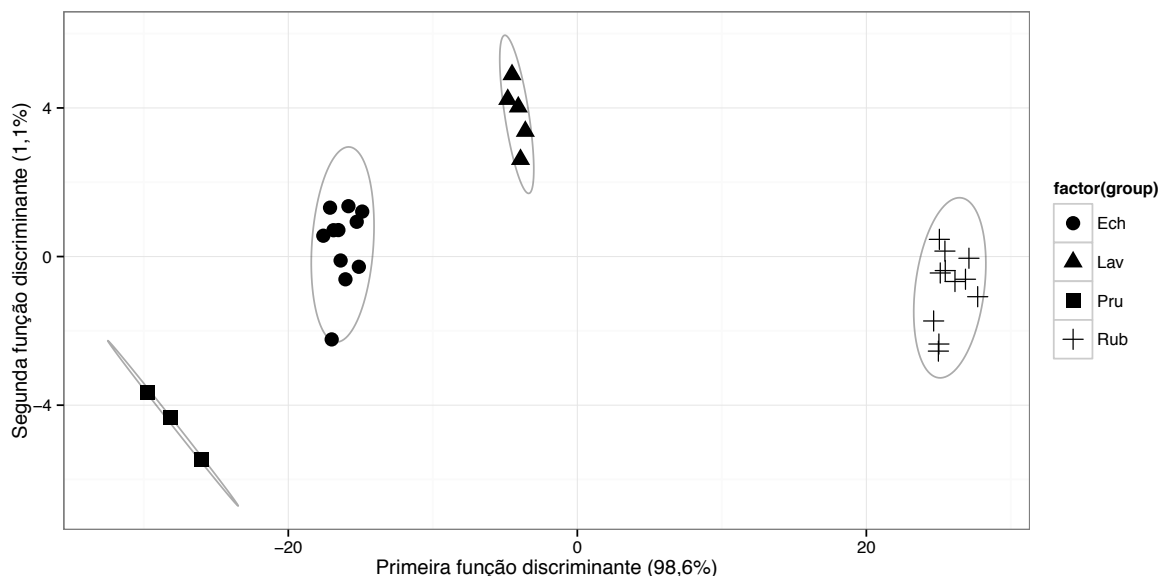


Figura 14 – Análise discriminante linear de 4 grupos de méis monoflorais de cor âmbar: *Echium* (Ech), *Lavandula* (Lav), *Prunus* (Pru) e *Rubus* (Rub). As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

3.4.3. Classificação de amostras de cor escura

Neste caso, o estudo incidiu em três grupos de méis monoflorais de cor escura com representação maioritária de *Castanea* (4 amostras), *Erica* (8 amostras) e *Rubus* (3 amostras). As amostras de méis monoflorais *Echium*, *Prunus* e *Trifolium* foram excluídas por só existir uma amostra de cada. Os resultados são apresentados na Tabela 15, cuja estrutura é semelhante à das anteriores.

O melhor modelo de discriminação foi obtido pelo algoritmo AS com um sub-conjunto de 7 sensores e com 100% de classificações corretas na validação cruzada LOO. Uma solução diferente foi obtida para os algoritmos GEN e MLR (chegaram ambos à mesma solução) para o mesmo sub-conjunto de sensores, tendo também 100% de classificações corretas na validação cruzada LOO, mas com o critério de qualidade ccr12 inferior ao do algoritmo AS. Por isso, o modelo definido pelo algoritmo AS foi o selecionado.

Este algoritmo mostrou maior robustez na seleção de sub-conjuntos de sensores pois apresenta 100% de classificações corretas na validação cruzada LOO para modelos com 7 a 10 sensores, obtendo-se posteriormente problemas de colinearidade por inclusão

de novos sensores. Situação semelhante foi obtida com os algoritmos GEN e MLR, mas no modelo com 8 sensores a percentagem das classificações corretas na validação cruzada LOO baixou para 97%.

Tabela 15 – Seleção de variáveis independentes através de 3 algoritmos para obter modelos de discriminação entre 3 grupos de mel monofloral de cor escura.

| Nº sensores | Sensores selecionados | ccr12 | %CC DO | %CC LOO |
|----------------------|--|------------------|------------|------------|
| Algoritmo AS | | | | |
| 2 | 38,39 | 0.5928385 | 67 | 47 |
| 3 | 3,11,23 | 0.7680201 | 80 | 67 |
| 4 | 9,10,20,32 | 0.8570319 | 93 | 67 |
| 5 | 2,26,32,37,38 | 0.9511542 | 80 | 67 |
| 6 | 2,4,14,26,28,38 | 0.9870934 | 100 | 80 |
| 7 | 12,15,18,27,29,33,39 | 0.9993452 | 100 | 100 |
| 8 | 12,15,17,18,27,29,33,39 | 0.9997313 | 100 | 100 |
| 9 | 12,15,17,18,26,27,29,33,39 | 0.9999793 | 100 | 100 |
| 10 | 1,6,11,14,15,23,24,37,39,40 | 0.9999992 | 100 | 100 |
| 11 | 1,2,4,14,15,17,20,25,35,38,40 | 1.0000000 | 100 | 27 * |
| 12 | 1,4,6,12,13,15,18,24,28,35,37,40 | 1.0000000 | 100 | 40 * |
| 13 | 2,15,16,18,19,24,25,30,31,32,36,37,39 | 1.0000012 | 100 | 47 * |
| 14 | 2,11,12,14,15,18,19,20,22,25,29,30,32,36 | 1.0000452 | 100 | 44 * |
| 15-20 | ** | | | |
| Algoritmo GEN | | | | |
| | | | | |
| 7 | 3,10,15,21,32,35,39 | 0.9957496 | 100 | 100 |
| 8 | 2,4,14,23,25,26,28,38 | 0.9996903 | 100 | 97 |
| 9 | 2,4,14,20,23,25,26,28,38 | 0.9999630 | 100 | 100 |
| 10 | 2,9,10,11,12,15,24,27,30,33 | 0.9999988 | 100 | 100 |
| 11 | 2,9,10,11,12,15,24,27,29,30,33 | 1.0000000 | 93 | 27 * |
| 12 | 3,4,8,9,11,19,24,26,29,31,33,36 | 1.0000000 | 93 | 27 * |
| 13 | 3,13,16,19,24,25,27,28,31,34,37,39,40 | 1.0000013 | 93 | 20 * |
| 14 | 3,8,9,13,18,19,24,25,27,28,34,37,39,40 | 1.0000387 | 93 | 53 * |
| 15-20 | ** | | | |
| Algoritmo MLR | | | | |
| | | | | |
| 6 | 1,2,29,32,37,38 | 0.9870934 | 93,3 | 73,3 |
| 7 = GEN | | | | |
| 8 = GEN | | | | |
| 9 | 2,4,14,20,23,25,26,28,38 | 0.9999988 | 100 | 100 |
| 10 | 2,7,10,13,21,23,26,31,38,40 | 0.9999988 | 93,3 | 27 * |
| 11 | 1,3,12,15,17,27,29,33,38,39,40 | 1.0000000 | 86,7 | 27 * |
| 12 | 3,5,8,9,15,22,26,28,30,34,37,39 | 1.0000000 | 100 | 60 * |
| 13 | 3,4,5,6,7,8,16,19,21,31,36,39,40 | 1.0000013 | 100 | 20 * |
| 14 | 6,8,16,19,20,24,28,29,30,31,32,37,38,40 | 1.0000387 | 100 | 33 * |
| 15-20 | ** | | | |

%CC DO – Percentagem de classificações corretas com os dados originais

%CC LOO – Percentagem de classificações corretas com validação cruzada LOO

*) Problemas de colinearidade verificadas na validação cruzada LOO

**) não foi capaz de encontrar qualquer subconjunto não-colineares confiável

Para o modelo selecionado obtiveram-se duas funções discriminantes que explicam 99,98% e 0,02% da variância total dos dados originais. A Figura 15, semelhante à anterior, mostra que a primeira função discriminante é suficiente para separar os 3 grupos de mel de

cor escura. Nesta dimensão, as amostras de *Erica* são as que se diferenciam mais ao nível do perfil de sinais da LE (conjunto de amostras no quadrante positivo). As amostras de *Castanea* e *Rubus*, presentes no quadrante negativo desta função discriminante, são as que mais se aproximam no perfil de sinais mas com diferenças que permitem uma separação efetiva entre grupos.

Nestes três estudos de classificação de amostras de méis monoflorais, o algoritmo AS foi sempre o que permitiu obter o modelo mais simples, com desempenho de 100% de classificações corretas das amostras obtidas pela metodologia da validação cruzada LOO, mostrando o seu bom desempenho na seleção dos sub-conjuntos de variáveis.

A principal diferença entre os algoritmos AS e GEN é que, enquanto o algoritmo AS cria uma nova solução modificando apenas uma solução com um movimento local, o GEN cria soluções através da combinação de duas soluções diferentes, por isso, a eficácia de cada um dependerá do problema e das condições definidas no algoritmo. Neste trabalho, as condições iniciais dos algoritmos foram fixas, com o objetivo de verificar qual o algoritmo que mostrava melhor resposta na seleção do óptimo local com dados analíticos de uma LE. Em relação ao algoritmo MLR, nenhuma variável é colocada em “standby” duas vezes, pelo que este não permite um novo estudo associado à introdução de uma variável já usada no modelo obtido, o que limita a confirmação de que solução final seja o óptimo local.

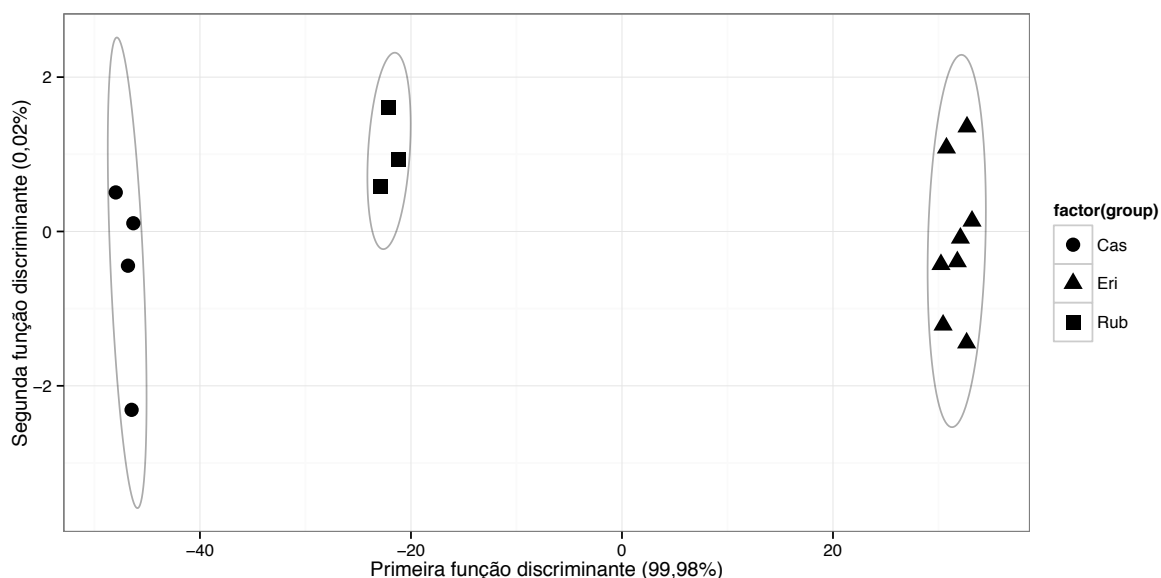


Figura 15 – Análise discriminante linear de 3 grupos de méis monoflorais de cor escura: *Castanea* (Cas), *Erica* (Eri) e *Rubus* (Rub). As elipses de cada grupo estão associadas ao intervalo de confiança a 95%, assumindo uma distribuição normal multivariada.

Globalmente, para os dados analíticos obtidos com a LE e seleção primária das amostras segundo três grupos de cor do mel, a aplicação do algoritmo AS na seleção dos melhores subconjuntos de variáveis para o modelo discriminante linear referentes a cada grupo de cor permite a classificação correta de todos os méis monoflorais. No caso do algoritmo GEN não se procurou otimizar as condições iniciais do algoritmo por ser um passo que levaria a várias tentativas e, por isso, a um processo demorado, não estando de acordo com o objetivo desejado neste trabalho.

Verificou-se que a LE construída foi eficaz na classificação do tipo de mel quando aliada à metodologia de tratamento de dados definida, atingindo-se um dos objetivos iniciais. Esta metodologia permitiu melhorar a eficácia da LE em relação ao trabalho realizado anteriormente pelo mesmo grupo de investigação [21]. A elevada variabilidade da composição química (ao nível dos açúcares maioritários), da cor e do perfil polínico existente nas amostras usadas neste estudo, que está associada a uma amostragem de 3 anos apícolas e que praticamente representa todas as regiões de Portugal continental, foi desejada para mostrar que a LE poderá ser uma ferramenta analítica eficaz a usar pelas empresas que produzem ou comercializam mel.

4. Conclusões

Neste trabalho aplicou-se uma LE potenciométrica construída pelo método *print-screen* e constituída por sensores de sensibilidade cruzada, com o objetivo principal de definir uma metodologia de tratamento de dados para discriminar diferentes tipos de méis monoflorais. Com vista a testar as capacidades analíticas da LE nos estudos do mel, obteve-se um conjunto numeroso de amostras com elevada variabilidade na sua composição polínica e química, que também foi visível através do amplo leque de cores (desde o extra-branco até ao escuro, correspondendo a variações de cor com valores entre 15,9 a 204,2 mm Pfund).

Foram usadas 89 amostras com perfis de pólen bastante distintos, dos quais 75 foram classificados em 8 tipos de mel monofloral diferentes. Os mais representados eram de méis monoflorais de *Lavandula sp.*, *Echium sp.*, *Rubus sp.* e *Erica sp.* (por ordem decrescente); com menor representação, méis monoflorais de *Prunus sp.*, *Castanea sp.*, *Trifolium sp.* e *Eucalyptus sp.* (os dois últimos, com 2 méis). A variabilidade das amostras é evidente ao nível do seu perfil polínico considerando quer o número de pólenes presentes quer a quantidade associada ao pólen predominante e dos restantes, aliás patente na variação da cor dentro de cada grupo de amostras de mel monofloral.

O método de HPLC usado para análise dos açúcares (glucose e frutose) das amostras mostrou ser preciso ($sr\% \leq 5\%$) e exato ($Er\% \leq 2\%$) em todos os estudos de verificação da qualidade analítica. Nas amostras, as concentrações de frutose variaram entre 33 e 57 g/100 g de mel e as de glucose entre 20 e 37 g/100 g de mel, correspondendo a níveis de monossacarídeos, calculada pela soma das concentrações da glucose e frutose, a variar entre 53,9 e 90,0 g/100 g de mel. O mel também pode sofrer adulterações com adição de adoçantes de baixo custo, tais como os xaropes naturais, por exemplo, os xaropes de alimentação de abelhas, produzido pela hidrólise do milho, da cana e açúcar de beterraba. Mas a adulteração de mel com açúcares ou xaropes não é facilmente detectada por análise direta de açúcares porque o produto adulterado também possui propriedades químicas e físicas semelhantes ao mel natural [54].

Nos dados obtidos com a LE, os perfis de sinais das amostras, mostraram uniformidade e sem presença de extremos que pudessem ser considerados *outliers*, indicativo da sua eficácia na análise das 89 amostras num dia. Estes dados correspondem a sinais de 40 sensores (dois sistemas de multi-sensores com 20 sensores iguais em cada). Devido à redundância da informação fornecida (em geral, os sinais mostram elevada colinearidade) foram testados três algoritmos meta-heurísticos de seleção de variáveis (arrefecimento simulado, genético e melhoria local) em paralelo com o método supervisionado ADL para estabelecer os modelos de classificação do tipo de mel.

Devido ao elevado número de tipos de mel presentes nas amostras foi necessário definir uma metodologia de tratamento de dados que permitisse efetuar a classificação das amostras, de forma robusta, no respetivo grupo de mel monofloral. Esta metodologia incidiu, como primeiro passo, em dividir as amostras em três grupos de cor redefinidas em função da escala Pfund: branco (inclui os méis extra-branco e branco), âmbar (inclui os méis âmbar extra claro, âmbar claro e âmbar) e escuro (constituído pelos méis escuros). A classificação das amostras através da análise multivariada ADL, usando os perfis de sinais obtidos com a LE e a classificação do tipo de mel monofloral estabelecida pela análise polínica, foi efetuada de forma independente em cada um dos três grupos referidos. Este procedimento, associado à seleção dos sub-conjuntos de sensores obtida pelo algoritmo AS (algoritmo que apresentou sempre os melhores resultados), permitiu obter o modelo de discriminação linear mais simples (com menor número de sensores no modelo ADL) com valor de ccr12 elevado e, principalmente, com 100% de classificações corretas com a validação cruzada LOO. Aliás, só é necessário usar a primeira função discriminante linear para garantir a previsão correta na classificação de méis monoflorais usando a LE potenciométrica com o procedimento descrito.

No estudo das amostras de cor branco obteve-se um modelo de ADL com 6 sensores que permitiu obter 100% de classificações corretas na validação cruzada LOO, entre as amostras dos dois grupos de *Echium* e *Lavandula*. Para as amostras de cor âmbar, o melhor modelo obtido continha 16 sensores e permitiu discriminar corretamente (100% de classificações corretas na validação cruzada LOO) 4 tipos de mel: *Echium*, *Lavandula*, *Prunus* e *Rubus*. No grupo de cor escura, o modelo obtido para a ADL envolvia 7 sensores selecionados e, tal como nos casos anteriores, permitiu discriminar com 100% de classificações corretas na validação cruzada LOO, 3 grupos de mel monofloral: *Castanea*, *Erica* e *Rubus*.

Os resultados obtidos neste trabalho permitiram mostrar as potencialidades da LE na análise de produtos alimentares, que aqui incidiu na análise qualitativa. Por exemplo, na continuidade deste trabalho, o próximo estudo será usar méis padrão monofloral (méis com elevada percentagem do pólen predominante) que permitirão calibrar a LE para posterior uso na classificação de novas amostras. Para tal será também necessário averiguar qual o melhor procedimento a usar no tratamento dos dados, de modo a obter metodologias de classificação robustas na previsão e permitir a análise rotineira de classificação dos méis. O desempenho analítico observado pela LE construída sugere que esta poderá ter um conjunto mais amplo de aplicações, além das referidas, como por exemplo, a análise quantitativa de compostos químicos maioritários do mel.

Referências

- [1] Honey Quality Standard, Codex Alimentarius, draft CL 1998/12-S
- [2] E. Crane, Bees and Beekeeping: Science, Practice and World Resources, Heinemann Newnes, Oxford, 1990
- [3] D. W. Ball, The chemical composition of honey, *Journal of Chemical Education* 84 (2007) 1643-1646
- [4] J.W. White, Honey, *Advances in Food Research* 22 (1978) 287-374
- [5] S. Bogdanov, Honey for nutrition and health: a review, *Journal of the American College of Nutrition* 27 (2008) 677-689
- [6] R.W. Thorp, The collection of pólen by bees, *Plant Systematics and Evolution* 222 (2000) 211-223
- [7] L.C. Marchini, A.C. Moreti, Comportamento de coleta de alimento por *Apis mellifera* L., 1758 (Hymenoptera, apidae) em cinco espécies de eucalyptus, *Archivos Latinoamericanos de Produccion Animal* 11 (2003) 75-79
- [8] V.M. Bryant, G.D. Jones, The R-values oh honey: pollen coefficients, *Palynology* 25 (2001) 11-28
- [9] J. Louveaux, A. Maurizio, G. Vorwohl, Methods of melissopalynology. *Bee World* 59 (1978) 139-157
- [10] W. Von Der Ohe, L.P. Oddo, M.L. Piana, M. Morlot, P. Martin, Harmonized methods of melissopalynology, *Apidologie* 35 (2004)
- [11] K. Ruoff, S. Bogdanov, Authenticity of honey and other bee products, *Apiacta* 38 (2004) 317-327
- [12] S.M.A. Pires, T. Rodrigues, A. Rocha, A. Pajuelo, O. Pereira, Caracterização polínica do mel de Trás-os-Montes e Alto Douro, *Revista Portuguesa de Zootecnia* 1 (2005) 87-99
- [13] M.G. Piazza, L.P. Oddo, Bibliographical review of the main European unifloral honeys, *Apidologie* 35 (2004) S94-S111
- [14] S. Bogdanov, Harmonised methods of the international honey commission, *International Honey Commission*, 2002
- [15] L.M. Estevinho, X. Feás, J.A. Seijas, M. Pilar Vázquez-Tato, Organic honey from Trás-Os-Montes region (Portugal): Chemical, palynological, microbiological and bioactive compounds characterization, *Food and Chemical Toxicology* 50 (2012) 258-264
- [16] L.R. Silva, R. Videira, A.P. Monteiro, P. Valentão, P.B. Andrade, Honey from Luso region (Portugal): Physicochemical characteristics and mineral contentes, *Microchemical Journal* 93 (2009) 73-77
- [17] F. Costa-Silva, M. Maia, C.C. Matos, E. Calçada, A.I.R.N.A. Barros, F.M. Nunes, Selenium content of Portuguese unifloral honeys, *Journal of Food Composition and Analysis* 24 (2011) 351-355
- [18] Programa Apícola Nacional 2014-2016, Gabinete de Planeamento e Políticas, Ministério da Agricultura do Mar, do Ambiente e do Ordenamento do Território, Abril 2013
- [19] YU. Vlasov, A. Legin, A. Rudnitskaya, C. DI Natale, A. D'Amico, Nonspecific sensor arrays ("electronic tongue") for chemical analysis of liquids, *IUPAC Technical Report, Pure and Applied Chemistry* 77 (2005) 1965-1983
- [20] K. Toko, Electronic sensing of tastes. *Electroanalysis* 10 (1998) 657
- [21] L.A. Dias, A.M. Peres, M. Vilas-Boas, M.A. Rocha, L. Estevinho, A.A.S.C. Machado, An electronic tongue for honey classification, *Microchimica Acta* 163 (2008) 97-102
- [22] C. Eckerta, C. Lutz, J. Breitzkreutz, K. Woertz, Quality control of oral herbal products by an electronic tongue—Case study on sage lozenges, *Sensors and Actuators B* 156 (2011) 204-212

- [23] J. Gallardo, S. Alegret, M. del Valle, Application of a potentiometric electronic tongue as a classification tool in food analysis, *Talanta* 66 (2005) 1303-1309
- [24] A. Legin, A. Rudnitskaya, L. Lvova, Yu. Vlasov, C. Di Natale, A. D'Amico, Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception, *Analytica Chimica Acta* 484 (2003) 33-44
- [25] Z. Wei, J. Wang, Classification of monofloral honeys by voltammetric electronic tongue with chemometrics method, *Electrochimica Acta* 56 (2011) 4907-4915
- [26] K. Tiwari, B. Tudu, R. Bandyopadhyay, A. Chatterjee, Identification of monofloral honey using voltammetric electronic tongue, *Journal of Food Engineering* 117 (2013) 205-210
- [27] Z. Wei, J. Wang, X. Zhang, Monitoring of quality and storage time of unsealed pasteurized milk by voltammetric electronic tongue, *Electrochimica Acta* 88 (2013) 231-239
- [28] S. Burattia, S. Benedettia, M. Scampicchioa, E.C. Pangerod, Characterization and classification of Italian Barbera wines by using an electronic nose and an amperometric electronic tongue, *Analytica Chimica Acta* 525 (2004) 133-139
- [29] P.A. Ulloa, R. Guerra, A.M. Cavaco, A.M.R. da Costa, A.C. Figueira, A. F. Brigas, Determination of the botanical origin of honey by sensor fusion of impedance e-tongue and optical spectroscopy, *Computers and Electronics in Agriculture* 94 (2013) 1-11
- [30] A. Rudnitskaya, L.M. Schmidtke, I. Delgadillo, A. Legin, G. Scollaryd, Study of the influence of micro-oxygenation and oak chip maceration on wine composition using an electronic tongue and chemical analysis, *Anal. Chim. Acta* 642 (2009) 235
- [31] R. Martínez-Máñez, J. Soto, E. Garcia-Breijo, L. Gil, J. Ibáñez, E. Llobet, An "electronic tongue" design for the qualitative analysis of natural waters, *Sensors and Actuators B* 104 (2005) 302-307
- [32] L. Escuder-Gilaberta, M. Peris, Review: Highlights in recent applications of electronic tongues in food analysis, *Analytica Chimica Acta* 665 (2010) 15-25
- [33] L.G. Dias, A.M. Peres, T.P. Barcelos, J. Sá Morais, A.A.S.C. Machado, Semi-quantitative and quantitative analysis of soft drinks using an electronic tongue, *Sensors & Actuators: B. Chemical* 154 (2011) 111-118
- [34] K. McGarigal, S. Cushman, S. Stafford, *Multivariate Statistics for Wildlife and Ecology Research*, Springer-Verlag New York, Inc, 2000
- [35] J.M. Graham, Problems with stepwise procedures in discriminant analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, LA (2001)
- [36] J. Cadima, J.O. Cerdeira, M. Minhoto, Computational aspects of algorithms for variable selection in the context of principal components, *Computational Statistics & Data Analysis* 47 (2004) 225-236
- [37] J. Cadima, J.O. Cerdeira, P.D. Silva, M. Minhoto, The subselect R package, 2012 (<http://cran.r-project.org/web/packages/subselect/vignettes/subselect.pdf>)
- [38] M. Gendreau, J. Potvin, *Handbook of Metaheuristics*, International Series in Operations Research & Management Science, V. 146, Springer, New York, 2010
- [39] K. Toko, Taste sensor, *Sensors and Actuators: B. Chemical* 64 (2000) 205-215
- [40] A.M. Peres, L.G. Dias, A.C.A. Veloso, S.G. Meirinho, J. Sá Morais, A.A.S.C. Machado, An electronic tongue for gliadins semi-quantitative detection in foodstuffs, *Talanta* 83 (2011) 857-864
- [41] K. Beullens, P. Mészáros, S. Vermeir, D. Kirsanov, A. Legin, S. Buysens, N. Cap, B. M. Nicolai, J. Lammertyn, Analysis of tomato taste using two types of electronic tongues, *Sensors and Actuators: B. Chemical* 131 (2008) 10-17

- [42] A. Legin, A. Rudnitskaya, L. Lvova, Y. Vlasov, C. Di Natale, A. D'Amico, Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception, *Analytica Chimica Acta* 484 (2003) 33-44
- [43] L.A. Dias, A.M. Peres, A.C.A. Veloso, F.S. Reis, M. Vilas-Boas, A.A.S.C. Machado, An electronic tongue taste evaluation: identification of goat milk adulteration with bovine milk, *Sensors & Actuators: B. Chemical* 136 (2009) 209-217
- [44] A. Legin, A. Rudnitskaya, Yu. Vlasov, C. Di Natale, E. Mazzone, A. D'Amico, Application of electronic tongue for qualitative and quantitative analysis of complex liquid media, *Sensors and Actuators: B. Chemical* 65 (2000) 232-234
- [45] N. Major, K. Markovic, M. Krpan, G. Saric, M. Hruskar, N. Vahci, Rapid honey characterization and botanical classification by an electronic tongue, *Talanta* 85 (2011) 569-574
- [46] Z. Wei, J. Wang, W. Liao, Technique potential for classification of honey by electronic tongue, *Journal of Food Engineering* 94 (2009) 260-266
- [47] I. Escriche, M. Kadar, E. Domenech, L. Gil-Sánchez, A potentiometric electronic tongue for the discrimination of honey according to the botanical origin. Comparison with traditional methodologies: Physicochemical parameters and volatile profile, *Journal of Food Engineering* 109 (2012) 449-456
- [48] I.C.F.R. Ferreira, E. Aires, J.C.M. Barreira, L.M. Estevinho, Antioxidant activity of Portuguese honey samples: Different contributions of the entire honey and phenolic extract, *Food Chemistry* 114 (2009) 1438-1443
- [49] USDA Agricultural Marketing Service, United States Standards for Grades of Extracted Honey, USDA, Washington DC, May 23, 1985
- [50] J. Louveaux, A. Maurizio, G. Vorwohl, Methods of melissopalynology, International Commission for Bee Botany of IUBS, *Bee World* 59 (1978) 139-157
- [51] Honey for Nutrition and Health: A Review, Stefan Bogdanov, *Journal of the American College of Nutrition*, Vol. 27, No. 6, 677-689 (2008)
- [52] T. Næs, B. Mevik, Understanding the collinearity problem in regression and discriminant analysis, *Journal of Chemometrics* 15 (2001) 413-426
- [53] K. Varmuza, P. Filzmoser, Introduction to multivariate statistical analysis in chemometrics, CRC Press (Taylor & Francis), Boca Raton, FL, USA, 2009
- [54] Honey Authenticity: a Review, Stefan Bogdanov and Peter Martin, Swiss Bee Research Centre (2002)